```
OM nucleic - nucleic search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             December 3, 2000, 17:37:21 ; Search time 23795.2 Seconds
                                                 (without alignments)
                                                 973.543 Million cell updates/sec

Title:              US-09-227-881-2
Perfect score:      5304
Sequence:           1 actcttgttcagttacctc............caggcacctctcagcacagc 5304

Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0

Searched:           1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters:          2067340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :

         GenEmbl:*
     1:  gb_bal:*
     2:  gb_ba2:*
     3:  gb_om:*
     4:  gb_ov:*
     5:  gb_pat:*
     6:  gb_ph:*
     7:  gb_pll:*
     8:  gb_pl2:*
     9:  gb_pr1:*
    10:  gb_pr2:*
    11:  gb_pr3:*
    12:  gb_ro:*
    13:  gb_sy:*
    14:  gb_un:*
    15:  gb_fun:*
    16:  gb_hum1:*
    17:  gb_hum2:*
    18:  gb_in:*
    19:  gb_om:*
    20:  gb_or:*
    21:  gb_ov:*
    22:  gb_pat:*
    23:  gb_ph:*
    24:  gb_pl:*
    25:  gb_ro:*
    26:  em_sts:*
    27:  em_sy:*
    28:  em_vi:*
    29:  em_vr:*
    30:  gb_ba3:*
    31:  gb_in1:*
    32:  gb_in2:*
    33:  gb_in3:*
    34:  gb_pl3:*
    35:  gb_pr4:*
    36:  em_bal:*
    37:  em_ba2:*
    38:  em_htg1:*
    39:  em_htg2:*
    40:  em_htg3:*
    41:  em_htg4:*
    42:  em_htg5:*
    43:  em_htg6:*
    44:  em_htg7:*
    45:  em_htg8:*
    46:  em_htg9:*
    47:  em_htg10:*
    48:  em_htg11:*
    49:  em_htg12:*
    50:  em_htg13:*
    51:  em_htg14:*
    52:  em_htg15:*
    53:  em_htg16:*
    54:  em_htg17:*
    55:  em_htg18:*
    56:  em_htg19:*
    57:  em_htg20:*
    58:  em_htg21:*
    59:  em_htg22:*
    60:  em_htg23:*
    61:  em_hum3:*
    62:  em_hum4:*
    63:  em_hum5:*
    64:  em_hum6:*
    65:  gb_pr5:*
    66:  gb_pr6:*
    67:  gb_pr7:*
    68:  gb_htg1:*
    69:  gb_htg2:*
    70:  gb_htg3:*
    71:  gb_htg4:*
    72:  gb_htg5:*
    73:  gb_htg6:*
    74:  gb_htg7:*
    75:  gb_htg8:*
    76:  gb_htg9:*
    77:  gb_htg10:*
    78:  gb_htg11:*
    79:  gb_htg12:*
    80:  gb_htg13:*
    81:  gb_htg14:*
    82:  gb_htg15:*
    83:  gb_htg16:*
    84:  gb_htg17:*
    85:  gb_htg18:*
    86:  gb_htg19:*
    87:  gb_htg20:*
    88:  gb_htg21:*
    89:  gb_htg22:*
    90:  gb_htg23:*
    91:  gb_sts1:*
    92:  gb_sts2:*
    93:  gb_vi1:*
    94:  gb_vi2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5274	99.4	79376	65	Z98750	Z98750 Human DNA s
2	5253.4	99.0	5300	11	AF007562	AF007562 Homo sapi
3	5223	98.5	170425	77	AC024490	AC024490 Homo sapi
4	1836	34.6	2800	66	HSMYOC1	AF049791 Homo sapi
5	417.8	7.9	1228	9	AB006686S1	AB006686 Homo sapi
6	407.8	7.7	1086	66	Z97171	Z97171 Homo sapien
7	189.8	3.6	161577	10	AC007688	AC007688 Homo sapi
8	189.8	3.6	193123	77	AC023790	AC023790 Homo sapi
9	189	3.6	199722	71	AC012404	AC012404 Homo sapi
10	187.8	3.5	97037	9	AC004973	AC004973 Homo sapi
11	187.8	3.5	135038	67	HUMWXD703	L78810 Homo sapien
12	185.8	3.5	76727	65	HS821D11	AL021453 Human DNA

ALIGNMENTS

```
c 13  185.8  3.5  152044  79  AC026395   AC026395   Homo sapi
c 14  185.8  3.5  157057  78  AC025947   AC025947   Homo sapi
c 15  185.8  3.5  161499  72  AC015488   AC015488   Homo sapi
  16  185.8  3.5  184656  86  AL161726   AL161726   Homo sapi
  17  185.8  3.5  200681  86  AC008755   AC008755   Homo sapi
  18  185.6  3.5  201372  86  AL157941   AL157941   Homo sapi
  19  184.8  3.5  146190  89  AP001813   AP001813   Homo sapi
  20  184.8  3.5  163494  89  AP002391   AP002391   Homo sapi
c 21  184.8  3.5  182241  89  AP001354   AP001354   Homo sapi
c 22  184.5  3.5  176029  11  AC011362   AC011362   Homo sapi
  23  184.2  3.5  130020  67  HUAC004525 AC004525   Homo sapi
  24  184    3.5  157304  78  AC024720   AC024720   Homo sapi
  25  184    3.5  187709  73  AC016168   AC016168   Homo sapi
  26  184    3.5  233734  70  AC011407   AC011407   Homo sapi
c 27  183.6  3.5  62070   88  AL358852   AL358852   Homo sapi
c 28  183.6  3.5  129370  89  AP000609   AP000609   Homo sapi
c 29  183.6  3.5  169333  89  AP002789   AP002789   Homo sapi
c 30  183.6  3.5  193171  74  AC018723   AC018723   Homo sapi
c 31  183.4  3.5  41407   11  AC053467   AC053467   Homo sapi
c 32  183.4  3.5  149138  79  AC026936   AC026936   Homo sapi
c 33  183.4  3.5  156331  90  HSJ19M11   AL121715   Homo sapi
  34  183.4  3.5  174974  87  AL356055   AL356055   Homo sapi
c 35  183.4  3.5  195364  65  HS431A14   HS431A14   Homo sapi
c 36  183.4  3.5  198583  74  AC019914   AC019114   Homo sapi
c 37  183.4  3.5  41936   35  CH192030879 AD000684  Homo sapi
c 38  183.2  3.5  70128   10  AC007536   AC007536   Homo sapi
c 39  183.2  3.5  119483   9  AC005588   AC005588   Homo sapi
c 40  183.2  3.5  145528   9  AC003665   AC003665   Homo sapi
  41  183.2  3.5  167943  65  HS267M20   HS267M20   Human DNA
c 42  183.2  3.5  170245  65  HS109F14   HS109F14   Homo sapi
c 43  183    3.5  242686   9  AL022721   AL022721   Human DNA
c 44  183    3.5  91767   10  AC000093   AC000093   Homo sapi
  45  183    3.5  112748  10  AC007242   AC007242   Homo sapi
```

RESULT   1

LOCUS   HS454G6/c   79376 bp   DNA   PRI   23-NOV-1999
DEFINITION Human DNA sequence from PAC 454G6 on chromosome 1q24. Contains
            trabecular meshwork inducible glucocorticoid response protein,
            TIGR, myocilin, ESTs and STS.
ACCESSION   Z98750
VERSION   Z98750.1 GI:2887277
KEYWORDS   1q24; myocilin; TIGR.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 79376)
  AUTHORS   Deadman,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-OCT-1997) Chromosome 1 Project Group
            (http://www.sanger.ac.uk/HGP/Chr1/) Sanger Centre, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT   On Feb 14, 1998 this sequence version replaced gi:2465060.
            IMPORTANT: This sequence is not the entire insert of clone 454G6.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variations annotated may not be found in the sequence submission
            corresponding to the overlapping clone as we submit sequences with
            only a small overlap as described above.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 1, constructed by the Sanger Centre chromosome 1
            mapping group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1/

This sequence has been finished according to sequence map criteria,
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The true left end of clone 454G6 is at 1 in this sequence. The true
left end of clone 560B9 is at 79273.
454G6 is from the library RPCI3 constructed at the Roswell Park
Cancer Institute by the group of Pieter de Jong.
For further details see http://bacpac.med.buffalo.edu/.

FEATURES   Location/Qualifiers
     source   1..79376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q24"
                     /clone="RP3-454G6"
                     /clone_lib="RPCI-3"
     repeat_region   435..472
                     /note="19 copies of 2 mer 82 % conserved"
     prim_transcript 1914..>3968
                     /note="match: multiple ESTs
                     match: R55676 AA043668 W63639 F12081 AA046699
                     match: F02925 AA131540 W00634 R36066 AA313383
                     match: AA163561 F02925 AA131540 W00634 R36066
                     match: AA313383 AA163561 N89173 AA174814 AA057059
                     match: AA329084 W47082 AA043955 AA341783 AA353681
                     match: AA046487 AA369741 H08313 AA186895 H32730
                     match: H08333 H08236 N42052 D61944 R27102 N32353
                     match: N30491 AA307150 AA192"
     repeat_region   3703..3746
                     /note="22 copies of 2 mer 89 % conserved"
     repeat_region   4051..4183
                     /note="AluSq repeat: matches 1..133 of consensus
                     incomplete repeat"
     repeat_region   4200..4502
                     /note="AluSq repeat: matches 2..301 of consensus"
     repeat_region   4659..4851
                     /note="AluSx repeat: matches 2..194 of consensus
                     incomplete repeat"
     repeat_region   5216..5345
                     /note="AluJo repeat: matches 132..1 of consensus
                     incomplete repeat"
     repeat_region   7759..7907
                     /note="MIR repeat: matches 174..1 of consensus"
     repeat_region   7933..9328
                     /note="TIGGER1 repeat: matches 1..1472 of consensus"
     repeat_region   9332..9626
                     /note="AluSg repeat: matches 1..289 of consensus"
     repeat_region   9639..10335
                     /note="TIGGER1 repeat: matches 1469..2174 of consensus"
     repeat_region   10343..10642
                     /note="AluSp repeat: matches 1..300 of consensus"
     repeat_region   10643..10856
                     /note="TIGGER1 repeat: matches 2175..2417 of consensus"
     misc_feature   <10903..>16855
                     /note="endogenous retroviral sequence"
     repeat_region   10949..11384
                     /note="LTR2 repeat: matches 31..449 of consensus"
     repeat_region   15835..15870
                     /note="18 copies of 2 mer 83 % conserved"
     repeat_region   16855..17286
                     /note="LTR2 repeat: matches 31..449 of consensus"
     repeat_region   17287..17574
                     /note="AluSg repeat: matches 15..300 of consensus"
     repeat_region   18294..18650
                     /note="THE1B repeat: matches 358..1 of consensus"
     repeat_region   18877..19180
                     /note="AluY repeat: matches 301..2 of consensus"
     repeat_region   19767..20013
                     /note="MIR2_internal repeat: matches 5002..4750 of
                     consensus"

AC000093   Homo sapi
AC007748 AC007748   Homo sapi
AC007242 AC007242   Homo sapi

```
repeat_region    20051..20118
                 /note="MLT1F repeat: matches 539..471 of consensus"
repeat_region    20130..20261
                 /note="AluJb repeat: matches 131..1 of consensus
                 Incomplete repeat"
repeat_region    20264..20722
repeat_region    20858..21223
                 /note="MLT1F repeat: matches 482..1 of consensus"
repeat_region    21216..21302
                 /note="MSTD repeat: matches 394..1 of consensus"
repeat_region    21403..21703
                 /note="MLT2_internal repeat: matches 4520..4433 of
                 consensus"
repeat_region    21978..22357
                 /note="MLT2_internal repeat: matches 3887..3580 of
                 consensus"
repeat_region    22363..22524
                 /note="MLT2_internal repeat: matches 3218..2839 of
                 consensus"
repeat_region    23286..>23680
                 /note="MLT2_internal repeat: matches 2495..2317 of
                 consensus"
misc_feature     22531..22839
                 /note="STS G07544"
misc_feature     complement(23652..24072)
                 /note="STS G07436"
repeat_region    23007..23309
                 /note="AluSx repeat: matches 302..1 of consensus"
repeat_region    23007..23309
                 /note="AluSp repeat: matches 3..301 of consensus"
repeat_region    24008..24290
repeat_region    25895..26364
                 /note="AluJo repeat: matches 298..6 of consensus"
repeat_region    26597..26697
                 /note="LTR2 repeat: matches 2..449 of consensus"
repeat_region    26719..27021
                 /note="AluSq repeat: matches 1..302 of consensus"
repeat_region    29037..29334
                 /note="AluSx repeat: matches 1..302 of consensus"
repeat_region    30028..30310
                 /note="AluSx repeat: matches 300..3 of consensus"
repeat_region    30402..30539
                 /note="AluSx repeat: matches 1..301 of consensus"
repeat_region    30402..30615
                 /note="MLT2B repeat: matches 264..404 of consensus"
repeat_region    30402..30615
                 /note="MLT2A repeat: matches 264..453 of consensus"
prim_transcript  <30801..>31136
                 /note="match: 5' EST AA256902 clone 682136"
repeat_region    30980..31222
                 /note="MER4B repeat: matches 199..431 of consensus"
repeat_region    31132..31528
                 /note="AluSc repeat: matches 299..1 of consensus"
repeat_region    31130..31654
                 /note="MER4B repeat: matches 417..540 of consensus"
repeat_region    31836..32135
                 /note="AluY repeat: matches 300..1 of consensus"
repeat_region    32200..32301
                 /note="AluJo repeat: matches 186..287 of consensus
                 Incomplete repeat"
repeat_region    32365..32493
repeat_region    33283..33399
                 /note="MIR repeat: matches 206..67 of consensus"
repeat_region    33408..33667
                 /note="MIR repeat: matches 77..196 of consensus"
repeat_region    33956..34043
                 /note="AluY repeat: matches 1..291 of consensus"
repeat_region    34725..34904
                 /note="MIR repeat: matches 80..167 of consensus"
repeat_region    34907..35207
                 /note="AluJo repeat: matches 299..127 of consensus
                 Incomplete repeat"
repeat_region    35212..35344
                 /note="AluY repeat: matches 300..1 of consensus"
repeat_region    35212..35344
                 /note="AluJb repeat: matches 133..1 of consensus

                 incomplete repeat"
repeat_region    36545..36842
                 /note="AluJo repeat: matches 299..1 of consensus"
repeat_region    38190..38379
                 /note="MIR3 repeat: matches 209..13 of consensus"
repeat_region    38182..38682
                 /note="AluSx repeat: matches 1..302 of consensus"
repeat_region    38836..38968
                 /note="MIR2 repeat: matches 145..2 of consensus"
repeat_region    38983..39279
                 /note="AluJb repeat: matches 3..296 of consensus"
repeat_region    39418..39773
                 /note="THE1A repeat: matches 354..2 of consensus"
repeat_region    39775..40120
                 /note="THE1B-INTERNAL repeat: matches 1580..1234 of
                 consensus"
repeat_region    40757..41023
                 /note="AluJo repeat: matches 37..301 of consensus
                 incomplete repeat"
repeat_region    41307..41589
                 /note="AluSq repeat: matches 21..303 of consensus


Query Match
Best Local Similarity 99.4%;           Score 5274;  DB 65;  Length 79376;
Matches 5293; Conservative 99.8%;  Pred. No. 0;  Mismatches 5;  Indels 6;  Gaps 1;
                                       0;  Mismatches

Qy      1 atctttgttcagtttaccttcagggcgtattatgaaatgaaatgaagatgataaccaaagtgaaag 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68406 ATCTTTGTTCAGTTTACCTTCAGGGCGTATTATGAAATGAAATGAAGATGATAACCAAAGTGAAAG 68347

Qy     61 tcccttataaactgtatagcccctcggcgtagtatctcttgggcaggagaggatggtaaaaggatca 120
          |||||||||||||||||||| |||||| ||| |||| |||||||||||| ||||||||||||||||
Db  68346 TCCCTTATAAACTGTATAGCCCTCCATTCGGATGTTAGTCTTTGGCAGGATGATAAAGAATCA 68287

Qy    121 ggaagaaggagtatccacgtagccaagtgtgtccaggcctgtctgctctgctctattctagtga 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68286 GGAAGAAGGAGTATCCACGTTAGCCAAGTGTGTCCAGGCCTGTGTCTGCTCTGCTCTTTTAGTGA 68227

Qy    181 cagatgttgtccctgacgagaagctctattcctcaggagaaacatcaacatccaatatggtaaatc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68226 CAGATGTTGCTCCCTGACAGAAGCTATTCTTCAGGAGAAACATCACAATATGGTAAATC 68167

Qy    241 catcaaacacaggagactaaagaaacagaaatgagagatgggagcttgcccaaggagaaaatgccag 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68166 CATCAAACACAGGAGACTAAGAACACAGGAATGAGATGAGGGCGCACTTGCCCAAGGAGAAAATGCCAG 68107

Qy    301 gagagcaaaataatgatgaaaaataaacttctccccttgttttaattcaggaaaaaaatg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68106 GAGAGCAAAATAATGATGAAAAAATAAACTTTTCCCCTTTGTTTTAATTCAGGAAAAAAATG 68047

Qy    361 atgaggaccaaaatcaatgaataaggagaaaacagcctcagaaaaaaagatgttcccaaattgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  68046 ATGAGGACCAAAATCAATGAATAAGGAGAAAAACAGCCTCAGAAAAAAAGATGTTTCCAAATTGG 67987

Qy    421 taattaagtatttgttccttgggagagacctccatgtgagcttgatggggaaaatgggaa 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67986 TAATTAAGTATTTGTTCCTTGGGAAGAGACCTCCATGTGAGCTTGATGGGGAAAATGGGAA 67927

Qy    481 aaacgtcaaaagcatcgatctgatcagcatcccaaagtggattattattttaaaacccagat 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67926 AAACGTCAAAAGCCATGATCTGATCAGCATCCCAAAGTGGATTATTATTTTAAAAACCAGAT 67867

Qy    541 ggcatcacctctggggagggcaagttcaggaggtcatgttagcaaagggacataacaataac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67866 GGCATCACTCTGGGGAGGGCAAGTTCAGGAGGTCAAGGTCATGTTAGCAAAGGACATAACAATAAC 67807

Qy    601 agcaaaaatcaaaattccgcaaaatgcaggagggaaaatgggggactggggaaagctttcataac 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67806 AGCAAAAATCAAAATTCCGCAAAATGCAGGAGGGAAAATGGGGACTGGGGAAAGCTTTCATAAC 67747

Qy    661 agtgattaggcagttgaccatgttcgcaacacctcccgtctataccagggaacacacaaaa 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  67746 AGTGATTAGGCAGTTGACCATGTTCCGCAACACCTCCCGTCTATACCAGGGAACACACAAAA 67687
```

```
Qy     721 attgactggcggctaagcctggactttccaagggaaatgaaaaactgagagagcaaaacaaa    780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67686 ATTTGACTGGGCTAAGCCTGGACTTTCAAGGGAAAATATGAAAAAACTGAGACAAAACAAAA  67627

Qy     781 gacatggttaaaggcaaccagaacattgtgagcctcaaggcagcagtgccctcagca        840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67626 GACATGGTTAAAAGGCAACCAGAACATTGTGAGCCTTCAAAGCAGCAGTGCCCTCAGCA  67567

Qy     841 gggaccctgaggcgcattggcctttagggaaggccagtgttctcaaggaatcttaagaaactc        900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67566 GGGAACCCTGAGGCGATTGGCCTTTAGGGAAGGCCAGTTTTCTTAGGAATCTTAAGAAACTC  67507

Qy     901 ttgaaagatcatgaatttaaccattttaagtatataaacacaaatatgccgatgatataatcag        960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67506 TTGAAAGATCATGAATTTTAACCATTTTAAGTATATAAAACAAAATATGCCGATGCATAATCAG  67447

Qy     961 tttagacatgggtcccaatttttatataagtctcaggccatacaagtaacgtgtcccagctcc      1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67446 TTTAGACATGGGTCCCAATTTTTATAAGTCTCAGGCCATACAAGTAACGTGTCCCAGCTCC  67387

Qy    1021 ggataggttcagaaatcattagaaatcactgttgtcccctacctacttttcagaaatgatc      1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67386 GGATAGGTTCAGAAATCATTAGAAAATCACTGTTGTCCCCATACCTAACTTTTCAGAAATGAATC  67327

Qy    1081 tgtgtcatagccccctccaacacagcccacagagctcgtgctcctactgtaccccaaa      1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67326 TGTGCAGCCCCATCCCGCCTCCCACGTGTCACTCTCGAGTCCCCAGCTACCCCAA  67267

Qy    1141 gtgtcagcccctctaacgtgttcatctcatgctcagtaggtggtcccattaccacactcacc      1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67266 TACAGCCAGAAGCTCCTGAGGTGGGTCTGTGTCTTACACCTACCTGTATGCTCTAC  67207

Qy    1201 tgtgctcagccctctgggccctccgcggtcacaggaagctccccccatactctctagacttcc      1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67206 TGTGCAGCCCATCCCGCTCCCACGGGGTCAAGCAATCTCTGTCTCAGCCTC  67147

Qy    1321 acctgagctcactgcaacctctgcctccaaggttcagcaatctcctgttctcaagcctcc      1380
           |||||||||||||||||||||||||||||||||||||||||||||
Db   67086 ACCTGAGCTCACTGCAACCTCTGCCTCCCAGGTCAATCTCACGTCCCCAGGTCAATCTCAGCCTCC  67027

Qy    1381 cgcgtagctgggactacaggcgcccgcgccaccgcggccggtaattttgttagtagagatggg      1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67026 CGCGTAGCTGGGACTACAGGCGCCCGCCACCGCGCCCGGCTAATTTTTGTTAGTAGAGATGGG  66967

Qy    1441 gtttcaccatattagcccggctagtctgctccatccctcgaacctcaggtgatcctccaacctc      1500
           |||||||||||||||||||||||||||||||||||||||||||||
Db   66966 GTTTCACCATATTAGCCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCACCTC  66907

Qy    1501 agccctcctaaagtgctgggattacaggcatgagcaaccaagggtcagtgt      1560
           |||||||||||||||||||||||||||||||||||||||||||
Db   66906 AGCCTCCTAAAGTGCTGGGATTACAGGCATGAGTCAGTGT  66847

Qy    1561 ttaataaggaataacttgaattgttactaacacagaggaaaagtcgtga      1620
           ||||||||||||||||||||||||||||||||||||
Db   66846 TTAATAAGGAAATAAACTTGAAATCGTTTACTAAACCAACAGACAAAGCTCTGA  66787

Qy    1621 taattccaggacttcctgggatgtgggaatggtgccatgagactgcctgcctagtcccagac      1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66786 TAAATTCAGGGACTTCTGGGAGTGTGCCCATGAGACTGCCCGCCTAGTCCCAGAC  66727

Qy    1681 cactggttcctccatccatccattttcaggctaagttaccacattttatt      1740
           ||||||||||||||||||||||||||||||||||||||||||||
Db   66726 CACTGGTTCCTCATCACTTTCTCCCTCATCCCCATTTCAGGCTAAGTTACCATTTTATT  66667

Qy    1741 caccatctgcttcttggttggtaagcctccacatcgttactgaaataagagtatacaataaactag      1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66666 CACCACTGCTTTTGTGGGTTAAGCCTTCCCACTTGTTACTGAAATAAGAGTATACAATAAAACTAG  66607
```

```
Qy    1801 tccaattggggccatctgtgtgtgtgtatagggagggagggcaatacccagagagact      1860
           ||||||||||||||||||||||||||||||||||||||||||||
Db   66606 TTCCATTTGGGGCCATTCTGGTGTGTGTGTGTGTATAGGGGGGAGGGGGCCATACCCCAGAGACTCCT  66547

Qy    1861 tgaagcccccgcggcagaggttctcctctccagctggggagccctgcaaggcaccggggtcc      1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66546 TGAAGCCCCCGGCGAGAGTTTCCTCTCCAGCTGGGGAGCCCTGCAAGGCACCCGGGGTCC  66487

Qy    1921 tgggtgtcctgagcaacctgccaacccgtccactgtcactggtgttgttattcactctctagg      1980
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   66486 TGGGTGTCCTGACCAACCTGCCAGCCCGTGCCACTGGTTGTTGTTTTGTTATTCACTCTCTAGG  66427

Qy    1981 gacctgttgcttctctattcctgtgactcgttcattcaatcaaggcattcattgacaatt      2040
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   66426 GACCTGTTGCCTTTCCTATTTCTAATTTCCTGTGACCTTCATTTCAATCAAGGCATTCATTGACAATT  66367

Qy    2041 tattgagtacttatatctgccacacaccaggagacacagaacaaatgtgtgatgcacactgc      2100
           |||||||||||||||||||||||||||||||||||||||||||||
Db   66366 TATTTGAGTACTTATATCTGCCAGACACAGAAGACACAGACAAATGTGTGATGCACACTGC  66307

Qy    2101 cctaccttcgtggagttgacagttcctggaacgttcaatcaaggttttgtctctgt      2160
           |||||||||||||||||||||||||||||||||||||||||||
Db   66306 CCTAACCTTCGTGGAGTTGACAGTTTCTCAGACACAGAGGACACAGACAAATGTGTCGC  66247

Qy    2161 gccaacttaaaccaggcccaacccctccctagcgccccccctgcacctcagaacctca      2220
           ||||||||||||||||||||||||||||||||||||||||||
Db   66246 GCCAACTTAAACCCAGTGCTGAAAGAAAGGGTCATGGAAGACAAAATGGTCGC  66187

Qy    2221 agctccttacaagggccaacctccctccacctagagccgccccctgcctccatcctgccggagg      2280
           ||||||||||||||||||||||||||||||||||||||||||||
Db   66186 AGCATCCCTTACAACGGGTCCTGAAATCGTCCTGGAATCCTGGAGACTCCTTGGGCTCCAGGCT  66127

Qy    2281 ccccgccaagccgcgagtcttccaagccctccaagctcagtcagctgcctgcctggcct      2340
           |||||||||||||||||||||||||||||||||||||||||||
Db   66126 CCCCCAACCCCAGTCTTCCAAGCTCCTCCATCAGTCAGTCACGCGGCCTGGCCT  66067

Qy    2341 gcctccgtcttcccgttgaatcgttcgtggttgcatctgagctcggttggctccaggct      2400
           |||||||||||||||||||||||||||||||||||||||||||
Db   66066 GCCTCGCTTCCCACGGTTGAATCGTTCCTGGTGTCAGGGAAAGGCTGAAGTCCTTGGGCTCCAGGCT  66007

Qy    2401 ccagaaaggaaatggagaaggggaaaactagtctaacggaaacagtgtttc      2460
           |||||||||||||||||||||||||||||||||||||||||||
Db   66006 CCAGAAAAAGGAAAATGGAGAAGGGGAAAACTAGTCTAACGGAAACAGTGTTTC  65947

Qy    2461 ctccagagggaaaggggccctccaacgtccagggaaggagaattccagtgcaggggg      2520
           ||||||||||||||||||||||||||||||||||||||||||
Db   65946 CTCAGAGGCAGAGGGAAAGGGGCCTCCACGTCCAGGGAAGGGGAGAGAATTCCAGGGG  65887

Qy    2521 tggggacgctggggctgaaagctgaaagacaggagaaaggggcctgaa      2580
           |||||||||||||||||||||||||||||||||||||||||
Db   65886 TGGGGACCGCTGGGGCTGAAGTCCTGAGCGGTGAAAAGGCCAGGAGAAAGGGGCTGAA  65827

Qy    2581 gctgcccagatgttcagtgttcagtgttgttccgctgagctgcttgctgtgagc      2640
           |||||||||||||||||||||||||||||||||||||||||
Db   65826 GCTGCCCAGATGTTCAGGTTGTTCACGGGCTTCCTTGCTGTGAGC  65767

Qy    2641 cttttctttctctctttttctctgcttgggagagaagagtcattcatgaaggatgcagtttc      2700
           |||||||||||||||||||||||||||||||||||||||
Db   65766 CTTTTTATCTTTATCTTGCCTTGGAGGGAAGGAGTCATTTCATGAAGGATGCAGTTTC  65707

Qy    2701 ataaagtcaagctgttaaaatccaggtgtcatggtgatgggacgttaaagccctttat      2760
           ||||||||||||||||||||||||||||||||||||||||
Db   65706 ATAAAGTCAGCTGTTAAAATTCCAGGTGTTTCCTTCACGAAGGCCTTTAT  65647

Qy    2761 ttaatgggaatatcaggaagcgagctcagttcatttcctaatttaattcacggaagaagtgac      2820
           ||||||||||||||||||||||||||||||||||||||
Db   65646 TTAAATGGGAAATATAGGAAGCGAGCTCATTTCCTTAAATTCACGGAAGAAGTGAC  65587

Qy    2821 tggagtctttctttcatgtctctttgggcaacactcagccctgtggttggacttggctta      2880
           ||||||||||||||||||||||||||||||||||||||
Db   65586 TGGAGTCTTTTCTTTCATGTCTTCTTCTGGGCAACACTCAGCCCTGTGGTGGACTTGGCTTA  65527

Qy    2881 tgcaagacggtcgaaacccttggaactcaggagactcggttctcttctggttctgccatt      2940
           ||||||||||||||||||||||||||||||||||
Db                                                                      ...
```

```
Db 65526 TGCAAGACGGTGGAAAACCTTGGAATCAGGAGACTCGGTTCTTTCTGGTTCTGCCATT 65467
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2941  ggttggctgttgccacctgggcaagtgtctctccttccctgggccatagttcttctgct 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65466 GGTTGGCTGTTGCCACCTGGGCAAGTGTCTCTCCTTCCCTGGGCCATAGTTCTTCTGCT 65407

Qy 3001  ataaagacccttgcactctcgtgttctgtgaacacttccctgtgatcctctgtgaggggg 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65406 ATAAAGACCCTTGCACTCTCGTGTTCTGTGAACACTTCCCTGTGATCCTCTGTGAGGGG 65347

Qy 3061  ggatgttgagagagggcagagagctgagcctgagctgagcagccacggggggaggg 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65346 GGATGTTGAGAGGGGGAAGGGAGAAGGAGCTGAGCCACAGGGGGAGGTGGAGG 65287

Qy 3121  ggacaggaaggagggcaggaggacaggcaggcagaaggcgggtgctccatcacagctgac 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65286 GGACAGGAAGGCAGGCAGAAGGCTCCATCACGTGGAAATGCCAGACTC 65227

Qy 3181  caggacggagagcccaagtgcctccaggaaagcttcaagttcaggaatgctcagggacca 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65226 CAGGACCGGAGAGCCCACAAATGCTTCAGGAAAGCTCAATGAACCCCACATTTTCCT 65167

Qy 3241  tcccctagacagacaatggcattggccaataaccaaaagcatgcaagaaaga 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65166 TCCCTAAGCATAGCACAATGGCATTTGCCAATAACCAAAAGCATGCAAGGAAAGA 65107

Qy 3301  ggtagctcttttgcctggttcattcaaaagtgggctctgttcaaggcaaggaatt 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65106 GGTAGCTTTTTGCCTGGCATTCAAAAGTGGGCTCTGTTCAAGGCAAGGGAATT 65047

Qy 3361  gttctaggagggcaggggctatatctgtggggaaaaatcagttcaagtttcaagt 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 65046 GTTCTAGGAGGGCAGGGGCTATATGTGGGGAAAAATCAGTTCAAGGAAGTCGGGAGA 64987

Qy 3481  acagattcattcaaggcgagtgggcagttctgggaatgtgatgttatgagcaagcaagaaa 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64926 ACAGATTCATTCAAGGCGAGTGGGCAGTTATGTCCACGTGATCCCTGG 64867

Qy 3601  cctgatttctaatactatattattttttcctttacaactgactgagtaattcacaag 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64866 GTTCTAGGAGGCAGGGGCTATATGTGGGGAAAAATCAGTTCAAGGAAGTCGGGACA 64807

Qy 3661  gtagtaactgagctgaagattacttactgagttcttccctattaggagaactctttctgt 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64806 CCTGATTTCTAATACTATATTTTCCTTTACAACTGACTGAGTAATTCACAAG 64747

Qy 3721  ggagttagcacgcacaagggcaatccgtttctttaaacggttaaacattcctaagag 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64746 GTAGTAACTGAGCTGAAGATTACTTAGTTCTCCCTTATTAGGAGAACTCTTTCTGT 64687

Qy 3781  ttaagccaaacagattcaagcctaggtcctgtactatatgattggtttttttttgaaaaat 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64686 GGAGTTAGCACACAAGGGCAAATCCGTTTCTTTAAACGGTTAAACATTCCTAAGAG 64627

Qy 3841  catttcagcgtatttacttacttctgattcagaaatgaacactagtaccctttggtcagctg 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64626 TAAAGCCAAACAGATTCAAGCCTAGGTCTGTACTATATGATTGGTTTTTTTGAAAAAT 64567

Qy 3901  taaacaaacacccctttgtaaaatgtctcaagttcaggtaactgcagaaccaatcaaat 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64566 CATTTCAGCGGATGTTTACTTACTTCTGATTCAGAAATGAACTAGTACCCTTTGGTCAGCTG 64507

Qy 3961  aagaatagaatctttagagcaaaactgtgttttctccactctgaggtgagtctgcaggg 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64506 TAAACAAACACCCCATTTGTAAAATGTCTCAAGTTCAGGTTAACTGCAGAACCAATCAAAT 64447
```

```
Db 64446 AAGAATAGAATCTTTAGAGCAAAACTGTGTTTTCTCCACTCTGAGGTGAGTCTGCAGGGC 64387
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4021  agtttggaaatatttaacttcacacaagtattgaacacctgtcttgttctggttattaacaacataaagt 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64386 AGTTTGGAAATATTTAACTTCACACAAGTATTGAACACTGTTGTTCTGGTTATTAACAACATAAAGT 64327

Qy 4081  tgctcaaaaggcaatcaatcaatcattatttcacaagtggcttaaagttacttccgaacagttcggtatat 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64326 TGCTCAAAAGGCAATCAATTATTTCAAGTGGCTTAAAGTTACTTCTGAACAGTTTGGTATAT 64267

Qy 4141  ttattgggctattgccatttgcttttgtctttttgggttttattcaataatgtaaagcag 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64266 TTATTGGGCTATTGCCATTTGCTTTTGTCTTTTTGGGTTTATTTAATAATGTAAAGCAG 64207

Qy 4261  tattttttaccacctctcactaactaattttacattttaacattttattccgaataggagccataaa 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64206 GGATTAATTAACCTAGACTTCAGAAAGCCTTGAATTGAGGAAAAAATTACATTTT 64147

Qy 4321  ctcaaagtcggtaataaagtaccctgtgatttctgattcaacaatagaaatcacagacatt 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64146 TGTTTTTACCACCTTCTAACTAAATTTAACATTTTATTCCGAATAGGAGCCATAAA 64087

Qy 4441  catgcacacacgagagtaagaacctacaggtaagaactgatttagaggctaacattga 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64026 GGAATTAATTAACCTAGACTTCAGAAAGCCTTGAATTGAGGAAAAAATTACATTTT 63967

Qy 4501  atttttgatattttgataatcatatttcattattttgttccttttgttaatctatcatttt 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 64086 CTCAAAGTCGGTTAATAAAGTACCCTGTGATTTTCGATTCAACATTCACAGACATT 64027

Qy 4561  atatactatattacagtttgttgcaggtacgttgaatgaaaatattatattctatctgg 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63906 ATTTTGATATTTTGATAATCATATTTCATTACACCTTGTTAATCTATATTT 63847

Qy 4621  catgcacacacgagagtaagaacctacaggtaagaactgatttagaggctaacattgga 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63846 ATATACTATATTACAGTTTGTTGCAGATTCCCCAGATTCACCAATGAGGTTCTTGG 63787

Qy 4681  tgcaagactgaaattaggaagaatacacacaccagtgttgtctttaaaagctagggtg 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63726 CATGCACACACGAGTAAGAACTTAGAGGCTAACATTGGTGCCTGAGA 63667

Qy 4741  agggggggaaatctgcagcttcctatatagaaagttgctctccccaaagatacacagttgttt 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63666 AGGGGGGAAATCTGCAGCTTCCTATATAGAAAGTTGCTCTCCCCAAAGATACACAGTTGTTT 63607

Qy 4801  cttgtgttctgggctggctggctgttatttttctctgttctcctgtacgtctcaaggacttgttt 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63606 CTTGTGTTCTGGGCTGGCTGTTATTTTTCTCTGTTCTCCTGTACGTCTCAAGGACTTGTTC 63547

Qy 4861  ggatcccacttcctagcatagtgcctggcacagttctcaaatggttctgcagag 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63546 GGATTCCCACTTCCTAGCATAGTGCCTGGCACAGTTCTCAAATGAGTTTGTTTGCAGAG 63487

Qy 4921  tgaatggaaatataaactagaaatatatatcctttgttgaatcagcacacagttctcaaatggttctgcagag 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63486 TGAATGGAAATATAAACTAGAAATATATCCTTTGTTGAATCAGCACACAGTTCTCAAATGGTTCTGCTGG 63427

Qy 4981  tgtaagttgtgtgtgtgtgtgtgtgtgtgtaaaaacaggtgga 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63426 TGTAAGTGTGTGTGTAC------GTGTGTGTGTGTGTGTGTGTAAAACCAGGTGGA 63373

Qy 5041  gatataggaactattattgggtatgggtgtccataaattgggatgtgttcttttttaaaaagaa 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63372 GATATAGGAACTATTATTGGGTATGGGGTCCATAAATTGGGATGTGTTCTTTTTTAAAAAGAA 63313
```

```
Qy  5101 actccaaacagacttcctggaaggttcattttcctaagaatcctctggcacgtgaaggcaa 5160
         |||||||||||||||||| |||||||||||| ||||| ||||| || |||||| |||||||||
Db 63312 ACTCCAAAACAGACTTCTGGAAGGTTATTTTCTTAAGAAATCTTGCTGGCAGGGTGAAGGCAA 63253

Qy  5161 ccccccctgcacagcgccccaccacccagccgcctcacgtgcctcctccccatgaag 5220
         |||||| |||||||||||||||| |||||||||||||||||||||||||||||||||| 
Db 63252 CCCCCCGTCCACAGCGCCCCACCGCCCAGCCGCCTCACGTGCCTCCTCCCCATGAAGG 63193

Qy  5221 gctggctcccagtatatataaaaccctctcgggggcatgggcccaagaggcccacc 5280
         |||||||||| |||||||||||||||||||| ||  || |||||| || |||||||
Db 63192 GCTGGCTCCCCAGTATATATAAAACCCTCTCGGAGCTCGGGCATGAGCCCAGCAAGGCCACC 63133

Qy  5281 catccaggcacctctcagcacagc 5304
         ||||||||||||||||||||||||
Db 63132 CATCCAGGCACCTCTCAGCACAGC 63109
```

RESULT 2

```
LOCUS       AF007562    5300 bp    DNA             PRI       18-MAR-1998
DEFINITION  Homo sapiens trabecular meshwork inducible glucocorticoid response
            protein (TIGR) gene, promoter region and partial mRNA sequence.
ACCESSION   AF007562
VERSION     AF007562.1  GI:2970123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5300)
  AUTHORS   Nguyen,T.D., Chen,P., Huang,W.D., Chen,H., Johnson,D. and
            Polansky,J.R.
  TITLE     Gene structure and properties of TIGR, an olfactomedin-related
            glycoprotein cloned from glucocorticoid-induced trabecular meshwork
            cells
  JOURNAL   J. Biol. Chem. 273 (11), 6341-6350 (1998)
  MEDLINE   98165818
REFERENCE   2  (bases 1 to 5300)
  AUTHORS   Nguyen,T.D., Chen,P., Chen,H. and Polansky,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1997) Ophthalmology, University of California San
            Francisco, 10 Kirkham Street, San Francisco, CA 94143-0730, USA
FEATURES             Location/Qualifiers
     source          1..5300
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1..>5300
                     /gene="TIGR"
     promoter        1..571
                     /gene="TIGR"
     mRNA            5272..>5300
                     /gene="TIGR"
                     /note="additional 3' sequence deposited as TIGR mRNA with
                     GenBank Accession Number U85257"
                     /product="trabecular meshwork inducible glucocorticoid
                     response protein"
BASE COUNT     1483 a    1152 c    1236 g    1429 t
ORIGIN
```

```
Query Match    99.0%;  Score 5253.4;  DB 11;  Length 5300;
Best Local Similarity    99.8%;  Pred. No. 0;
Matches 5293;  Conservative  0;  Mismatches  6;  Indels  6;  Gaps  3;
```

```
Qy    1 atctttgttcagtttaccccaaaggctattatgatgaatgtgagataaccaatgtgaaag 60
        |||||||||||||||||| ||| ||| ||||||||| ||||||||||  |||||||||||||
Db    1 ATCTTTGTTCAGTTTACCTCAGGGCCTATTATGAAATGAGATGAAACCAATGTGAAAG 60

Qy   61 tcctatataactgtatagcccccatccggatgatgtgataaatctccggcctgaaaaatca 120
        |||||||  ||||||||||| |||| |||||||||||  |   ||||  |||| |||||||||
Db   61 TCCTATAAAACTGTATAGCCTCCATTCGGATGATGTAGTTCTTTGGCAGGATGATAAAGAATCA 120

Qy  121 ggaagaaggagtatccacgttagccagtgtccaagtgtcctgtctgtcctcttattagtga 180
        |||||||||||||||  ||||||||||| |||| || |||| |||||||||  ||||| ||||
Db  121 GGAAGAAGGAGTATCCCACGTTAGCCAAGTGTCCAGGCTGTCGTGTCTGTCCTTATTTTAGTGA 180

Qy  181 cagatgttgctccctgacgagagctattccttcaggaaacatcaccatccaaatatggtaaatc 240
        ||||||||||||| |||||||||||||||||||||||||| ||||| ||||||||||| |||||
Db  181 CAGATGTTGCTCCCTGACGAGAGCTATTCTTCAGGAAACATCACCATCCAAATATGGTAAATC 240

Qy  241 catcaaacaggagctaagaaacacaggaatgagatgagaggaaaagtgccag 300
        |||||||||||| ||||||||||||||||| ||||||||||||||||||||||
Db  241 CATCAAACAGGAGCTAAGAAACACAGGAATGAGATGGGGGCACTTGCCCAAGGAAAAATGCCAG 300

Qy  301 gagagcaaataatgatgaaaaaataaactttccctttgtttttaatttcaggaaaaatg 360
        |||||||||||||| |||||||||||||||||||| |||||||| ||||||||||||||||
Db  301 GAGAGCAAAATAATGATGAAAAAATAAACTTTTCCCTTTGTTTTTAATTTCAGGAAAATG 360

Qy  361 atgaggaccaaaatcaatgaataataaggagaaaacagctcagaagaaaaagatgttccaaaattgg 420
        ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  361 ATGAGGACCAAAATCAATGAATAATAAGGAGAAAACAGCTCAGAAGAAAAAGATGTTCCAAAATTGG 420

Qy  421 taattaagtattgtccttgggaagagaacctccatgtgagcttgatgtgggaaaatgggaa 480
        |||||||| ||||| ||| |||||||| |||||| |||||||||||  ||||||||||| ||
Db  421 TAATTAAGTATTGTTCCTTGGGAGGGCAACTTCACGTGAGCTTAGCAAAAGGACATAACAATAAC 480

Qy  481 aaacgtcaaaagccatgaatctgatcagaatcccaaagtcccaaaaaggatataattaaaaaaccagat 540
        |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  481 AGCAAAATCAAAATCAAAATTCCGCAAATGCAGGAGGAAAATGGGGACTGGGAAGCTTTCATAAC 540

Qy  541 ggcatcaactctggggaggcaagttcagttggcgaaagggacaataacaatac 600
        |||||||||||| ||||||||||| ||||||||| ||||||||| |||||||||
Db  541 GGCATCACTCTCTGGGGAGGCAACTTCACGTCATGTTAGCAAGACATAACAATAAC 600

Qy  601 agcaaaatcaaaattccgcaaatgcaggaggaaaatggggactgggaaagctttcataac 660
        |||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||
Db  601 AGCAAAATCAAAATTCCGCAAATGCAGGAGGAAAATGGGGACTGGGAAGCTTTCATAAC 660

Qy  661 agtgattaaggcagttaaccatgttcggcaacacctcccgtctatacaggggaacacagca 720
        |||||||||||||||||  |||||||| |||||| |||| |||||||||||| ||||| |||
Db  661 AGTGATTAGGGCAGTTGACGTTCGCCAAGTCCCCCGTCTATACCAGGAACACAAAA 720

Qy  721 attgactgggctaagcctggacttcaagggaaatatgaaaaactgagagacaaaacaaa 780
        |||||||||||||||||||||||||| |||||||  |||||||||| || |||||||||
Db  721 ATTGACTGGGCTAAGCCTGGACTTTCAAGGGAAATATGAAAAACTGAGAGACAAAACAAA 780

Qy  781 gacatggttaaaggcaaccagaacattgtgagccttcaagcagcagagatcggccctcagca 840
        |||||||||||||||||  |||||||||||||||||||| ||||||||| |||| ||||||
Db  781 GACATGGTTAAAAGGCAACCAGAACAATGTGAGCCTTCAAAGCAGCAGTGCGCCCCTCAGCA 840

Qy  841 gggaccctgaggccattgccctttaggaaggccagtgttcttaaggaatcttaagaacc 900
        ||||||||||||||||||||||| ||||||| |||||| ||| ||||| ||| |||||||
Db  841 GGGACCCTGAGGCCATTTGCCTTTAGGAAGGCCAGTTCCTTTAAGGAATCTTAAGAAACTC 900

Qy  901 ttgaaagatcatgaatttaaccattttaaccataatatgggagtctaaggatgctatcag 960
        |||||||||||||||||||||||| |||||||| | |||| ||||||| | |||| ||||
Db  901 TTGAAAGATCATGAATTTAACCATTTTAAGTATAAAACAAAATATGGGAATCATAATCAG 960

Qy  961 tttagacatgggtcccaaatttatataagtcaggcataccaagaatacgtgtccccgctcc 1020
        |||||||||||||||||||| |||| |||||||| |||| |||| |||||||| |||||||
Db  961 TTTAGACATGGGTCCCAAATTTATATAAGTCAGGCATACCAAGATAACGTGTCCCAGCTCC 1020

Qy 1021 ggataggtcagaaatcattagaaatacactgtgtcccccatcctaacttctttcagaatgatc 1080
        |||||||||||||||||||||||||||| |||||||||||||| |||| |||||||||||||
Db 1021 GGATAGGTCAGAAATCATTAGAAATCACTTAGAAATCACTTAGAAATCACTGTCCCAATTAGAATC 1080

Qy 1081 tgtcatagccctcacacacggcccgatgtgtctgacctacacaccatccaacccaa 1140
        |||||||||| ||| |||||| |||  ||||||||| |||| |||| ||||| |||||
Db 1081 TGTCATAGCCCTCACACACAGCGCCCCAGTGTCTGACCTACACAACCTACAACCCAA 1140

Qy 1141 gtgcctcaaccattgtttaacgtgtcatctccagtaggtccattacaaaatgccacctccc 1200
        ||||||||||| |||||| ||||| || || ||||| ||| || ||| |||| |||||||||
Db 1141 GTGCCTCAACCATTGTTTAACGTGTCATCTCAGTAGGTCCATTACAAAATGCCACCTCCC 1200

Qy 1201 tgtgtcagcccatcccgctccacaggaagtctcccactctgcatcacagatgt 1260
        ||||||||||||||||||||||||
Db 1201 TGTGTCAGCCCATCCCGCTCCCC 1260
```

Db 1201 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1260
Db 1201 TGTGCAGCCAGGAAGCTCCGTGGAGGTGAGGTCTGTTGTTCTTACCACTTGATGTCTCTAC 1260

Qy 1261 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1320
Db 1261 TACAGCCAGGAAGCTCCGTGAGGGTGAGGGTCTGTGTGTCTTACACCTACCTGTATGCTTAC 1320

Qy 1321 acctgagctcactgaacctctgcctcccaaggttcaagcaattctcctgtctccagcctcc 1380
Db 1321 ACCTGAGCTCACTGCAACCTCTGCCTCCCCAGGTTCAAGCCAATTCTCCTGTCTCCAGCCTCC 1380

Qy 1381 cgggtagctgggactacaggcgcacgccaccgggcctaattcttgtattgtagtaggaagtggg 1440
Db 1381 CGGGTAGCTGGGACTACAGGCGCCCGCCACCGGGCCTAATTTTTGTATTTGTTAGTAGGAGGG 1440

Qy 1441 gtttcaccacatattagcccacctattgatctctgaagccccgggtcttgtcttgaactctgg 1500
Db 1441 GTTTCACCACATATTAGCCGCCTCCCCAGGTCTTGAACTCCTGGGCTCAGGCGATCCACCACCTC 1500

Qy 1501 agcctcctaaagtgctgggattacaggcatgagtcaccgcgcccggccaaggggtcagtgt 1560
Db 1501 AGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAAGGGGTCACTGT 1560

Qy 1561 ttaataaggaataacttgaatggttacctaaaacacaggaaacacagacaaaagctgtga 1620
Db 1561 TTAATAAGGAATAACTTGAATGGTTACTAAAACCAAGGAAACAGACAAAAGCTGTGA 1620

Qy 1621 taattcaggagttcttggatatgatggatctgaagagtggcatggaaacctgcagac 1680
Db 1621 TAAATTTCAGGGATTCTTGGGACGGGAAATGGTCCCATGACTCCGCTAGGTCAGGCGACTCCT 1680

Qy 1681 cactggtcctcatcactgttcttctttcccctcaatcccattcttcaggctaagttaccattttatt 1740
Db 1681 CACTGTTCCTCATCACTTTCTTTCTTCCCTCCATTTTCAGGCTAAGTTACCATTTTATT 1740

Qy 1741 caccactgcttctgttctgagaccctgccactgcttactctgaaaatataataaaactag 1800
Db 1741 CACCAGCTTTTGTGGGTAAGGCTCCACATCGTTACTGAAATAAGAGTATACATAAACTAG 1800

Qy 1801 ttccattttgggttgtcctgtgtatgtggcatgccctagtcccagac 1860
Db 1801 TTCCATTTGGGGCCATCTGTGTGTGTATAGGGGAGGGGATACCCCAGAGACTCCT 1860

Qy 1861 tgaagccccggcagagagtttcctctccagctgggggggagccctgcaagccagttggtcc 1920
Db 1861 TGAAGCCCCGGCGAGAGGTTTCCTCTCCAGCTGGGGGGAGCCCTGCAAGCCAGTTGGTCC 1920

Qy 1921 tgggtgtgtcctgagcaacctgccagccctgccactgcttactctgtctctcctctctgg 1980
Db 1921 TGGGGTGTCCTGAGCAACCTGCCAGCCCTGCCACTGGTTGTTTATCACTCTCTAGG 1980

Qy 1981 gacctcgttgcttctcattcctgtgtgactcgttcatccaggcattcattgacaatt 2040
Db 1981 GACCTCTGTTGCTTTCTATTTCTGTGTGACTCGTTCATCCAGGCATTCATTGACAATT 2040

Qy 2041 tattgagtacttatatctgccagacacagagagacaaaatgttgagcaagcagtcactgc 2100
Db 2041 TATTGAGTACTTATATCTGCCAGACACAGAGACAAAATGGTGAGCAAAAGCAGTCACTGC 2100

Qy 2101 cctacccttcgtggagttgacagtttctcattggaagagaaattaatagcca 2160
Db 2101 CCTACCTTCGTGGAGTTGACAGTTTCTCATGGAAGAGAAAATTAATAGCCA 2160

Qy 2161 gccaacttaacccagtgctgaaagaaagtaacacacctgaaagaattgtcgc 2220
Db 2161 GCCAACTTAAACCCAGTGCTGAAAGAATAAACACACTTGAAGAATTGTCGC 2220

Qy 2221 agcatcccttaacaaggccaactcctagcgccccctgctcctcatcgtcgtgccggagg 2280
Db 2221 AGCATCCCTTAACAAGGCCACTCCTTAGCGCCCCCTGCTCCTCATCGTCCCGGAGG 2280

Qy 2281 cccccaagccagtcttcctcccatcagtcacagcggtgcagctggct 2340
Db 2281 CCCCCAAGCCCGAGTCTTCCAAGCCCTCCTCCAGTCACGCGCTGGCCT 2340

Db 2281 CCCCCAAGCCCGAGTCTTCCAAGCCCTCCTCCAGTCACGCGCTGGCCT 2340

Qy 2341 gcctcgcttcccgtgaatcgttcctggcatctgagctggagactccttggccaggct 2400
Db 2341 GCCTCGCGTTCCCGTGAATCGTTCCTGGCATCTGAGCTGGAGACTCCTTGGCTCCAGGCT 2400

Qy 2401 ccagaaggaaatgggaacggggaaactgttctaacggggacagttgttc 2460
Db 2401 CCAGAAAAGGAAATGGAGAGAGGGAAACTAGTCTAACGGAGGAAGTCGGGAACAGTGTTTC 2460

Qy 2461 ctccagaggggaaagggggcctccacgtccaggaggacgaattccaaggaggagtggtgggagg 2520
Db 2461 CTCAGAGGGGAAAAGGGGCCTCCACGTCCAGGAGGAATTCCAGGAGGTGGGGGACTCCAGGAG 2520

Qy 2521 tggggacggctggggtgacgggtcgagccagaaggaggaaagttgttggaaagctgaa 2580
Db 2521 TGGGGACGGCTGGGGCTGAGGGGGTCGAGCCAGAAGGCAGAAAGGGTGAAAAGGCTGAA 2580

Qy 2581 gctgcccagagtgtcctgggcctgggcagtgggtgttcacggcagttttccttcacgagctgagc 2640
Db 2581 GCTGCCCAGATGTTCAGTGTTGGGCCTGGGCAGTGGGGTTTCACGGCAGTTTTCCTTCCTGTGAGC 2640

Qy 2641 ctcagaggggaaagggggctccacgtccaggaggagaattccaggaggtggggactccaggag 2700
Db 2641 CTCAGAGGGGAAAAGGGGGCTCCACGTCCAGGAGGAGAATTCCAGGAGGTGGGGGACTCCAGGAG 2700

Qy 2701 ataaagtcaagctgttaaaaaattccaagggtgtgtgcatgggttcttccttcccaagt 2760
Db 2701 ATAAAGTCAAGCTGTTAAAAAATTCCAGGGTGCATGGGTTTTCCTTCACGAAGGCCTTAT 2760

Qy 2761 ttaatggaaatataggagagcggcagcggcagtcattccctaaggccgttaattcacggaaggaagtgac 2820
Db 2761 TTAATGGGAATATAGGAAACGGACGGCGTCATTTCCTAGGCCCGTTAATTCACCGGAAGTGAC 2820

Qy 2821 tggagtctttctttctttcatgtctcttctgtgtgaacatactacctgactcagggaggctagg 2880
Db 2821 TGGAGTCTTTTTCTTTCTTTCATGTCTCTTCTGTGTGAACATACTACCTGACTCAGTTTA 2880

Qy 2881 tgcaagacggtcgaaaacctggaatcaggagcagagatcgggggattcttctggttctcttctctgccatt 2940
Db 2881 TGCAAGACGGTCGAAAACCTGGAATCAGGAGACTCGGGGTTTTCTTCTCGTTCTCTGCCATT 2940

Qy 2941 ggttggctcgtgtgcgcagccttggcaagtgctcattttcatcctgacagtagttctctgt 3000
Db 2941 GGTTGGCTCGTGCGCAGCCCTGGGCCATAGTCTTCTCTGCT 3000

Qy 3001 ataaagaccctctgcagctcctcgtgttctctgtgaacacttgattctctgaggggg 3060
Db 3001 ATAAAGACCCCTTGCAGCTCGTGTTTCTCGTGATTCTCTGAGGGGG 3060

Qy 3061 ggatgttgagaggggaaggcagagcagaagacggggaggg 3120
Db 3061 GGATTGTTGAGAGGGGAGAGGGGAAGGACCAGAGCTGAGGAGG 3120

Qy 3121 ggtcaggggaaggcggggaagtgtcttcctccacactgagacagactc 3180
Db 3121 GGACAGGAGAGGCAGCAGAAGCCTCCATCACAGTCGTGGCCCTTGCT 3180

Qy 3181 caggaccgagagccacaatgcttcagggaaagctcaatgaaccacgccacatttcct 3240
Db 3181 CAGGACCAGAGCCACACATGCTTCAGGAAAGCTCAATGAACCCAAACAGCACATTTTTCT 3240

Qy 3241 tccctaagctagacaatggcatttgccaataccaaaaaagaatgcagggt 3300
Db 3241 TCCCTAAGGCATAGACAATGGCCAATACCAAAAAAGAATGCAGAGACTGGT 3300

Qy 3301 ggtagcttttgtctggcattcaaaaaacttgggccaaccgagagaaatg 3360
Db 3301 GGTAGCTTTTTGCCTGGCATTCAAAAAACTGGGCCAAGAGTGGAAAATGCCAGAGATG 3360

Qy 3361 ttaaactttcaccctgaccagcacccacggcagctcagcagctgacacg 3420
Db 3361 TTAAACTTTTCACCCTGACCAGCACCCACGCAGCTCAGCAGCTGACTCGCACG 3420

```
Qy   3421 agtgacctgcacagccgcacgggggaggaggaagaagaaagaaagatagtgtgtgagcaagaaag 3480
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 AGTGACCTGCACAGCCGCAGGGCGAGGGAGGAGAAGAAAGAGAGGGATAGTGTATGAGCAAGAAAG 3480

Qy   3481 acagatccattcaagggcagtgggaattgacaccacgtggaattgaccatccactgtgatcctcg 3540
          ||||||||||||| || || |||||||||||||||||||||||||||||||||||||||||| ||
Db   3481 ACAGATTCATTCAAGGGCAGTGGGAATTGACCACCACGTGGAATTGAGTTAGTCCACGTGATCCTGG 3540

Qy   3541 gtcctaggggggcaggggctatatatgtggggggggaaatcagtccaaggctaacattggtcctgag 3600
          ||||||| |||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db   3541 GTTCTAGGAGGGCAGGGGCTATATTGTGGGGGAAAAATCAGTTCAAGGGAAGTCGGGAGA 3600

Qy   3601 cctggattctaatactatattttttttccttacacaagtgtaatctcgtagggacgcacaag 3660
          |||||| |||||||||| |||| ||||||||||||||||||||||||||||||||||||||
Db   3601 CCTGGATTTCTAATACTATATTTTTTCCTTTACACAGGTGTAATTCTGACGAAGTCACAAG 3660

Qy   3661 gtagtaactgagggctgtaagattactaactttttttCCTTATTAGGGAACTCTTTTTTCTCTTGT 3720
          ||||||||||| |||||||||| ||||||||| |||||||||||||||||||||||||||||
Db   3661 GTAGTAACTGAGGGCTGTAAGATTACTAGTTTCTCCTTATTAGGGAACTCTTTTTTCTCTGT 3720

Qy   3721 ggagttagcagcacaagggcaaatcccgttcctcttttctttctctcttaaggacttgtt 3780
          |||||||||||||| |||||||||||||| |||||||||||||||||||||||| ||||
Db   3721 GGAGCTTAGCAGCACAAGGGCAATCCCGTTCCTTTTAACAGGAGAAAAACATTCCTAAGAG 3780

Qy   3781 taaagccaaacagattcaagccctaggtcttgctcgactattgattggttctttttgaaaat 3840
          ||||||||||||||||||| |||| |||||||| ||||| || | |||||||||| ||||||
Db   3781 TAAAGCCAAACAGATTCAAGCCTAGGTCTTGCTCGACTATATGATTGGTTTTTTTGAAAAAT 3840

Qy   3841 cattcccgatgtgtttttactatctgattccagaaaatgaactgagctgtaccctttgtgtcagctg 3900
          |||||||||| || ||| ||||||||||||||| ||||||||||||||||||||||||||
Db   3841 CATTTCAGGGATGGTTTTACTTACTTGAGACAAAACTGTTGTATCTGATTCAGAAAAATGAGAG 3900

Qy   3901 taaacaaacaccccattttgtaaagttcaagttcaagtgctcaagttcaactgcagcaaccaatcaaat 3960
          ||||||||||| ||||||||||||||||||||| |||||||||||||||||||||||
Db   3901 TAAACAAACACCAGTTGTAAAGTCTCAAGTACAAGGCTTAACTGCACGAACCAATCAAA- 3959

Qy   3961 aagaatagaatcttagagcaacaactgttgtctctccac-cctggaaggtgagtctgccagg 4019
          ||||||||||| ||||||||||||||||||||||||||  |||||||||||||||||||||
Db   3960 AAGAATAGAATCTTTAGAGCAAAACTGTTGTTTCTCCACATCTGGAAGGTGAGTCTGCCAGGG 4019

Qy   4020 cagtttggaaatatcttagacagtgttactcacaagtattgccacagtagttaataacgtagtac 4079
          ||||||| |||||| |||||| |||||| |||||||||||| |||||||||||||||||||
Db   4020 CAGTTTGGAAATAATCTTAGACGTATTGACACTGTGTTGGTATTAAACCATAAAG 4079

Qy   4080 ttgctcaaggcaatcatcattcatttctttactttcacatgtctaaagttacttctgacagtttttggta 4139
          ||||||| |||||||||||||||||| ||||||||||||||||||||| ||| ||||||||
Db   4080 TTGCTCAAAGGCAATCATTCATTTTCAAGTTACTTCTGACAGTTTTGGTATA 4139

Qy   4140 tttattggtattgccaatttttgctttgtttattaatgtaaagca 4199
          ||||||||||| ||||||| |||||||||||||| |||||||||||
Db   4140 TTTATTGGCTATTGCCATTTTGCTTTTGTTTTCTCTTTGGGTTTATTGTAAAGCA 4199

Qy   4200 gggattattaacctacagtccagaaaagcctgtgaattgaatgagggaaaaaaatacgttt 4259
          ||||||||||||||||||||| |||| |||||||| ||||||||||||||||||||||
Db   4200 GGGATTATTAACCTACAGTCCAGAAAAGCCTGTGAATTTGAATGAGGAAAAAAAATTACATTT 4259

Qy   4260 ttatcttttaccacttctaactaaattttatcccaatagagccataa 4319
          || ||||||||||||||||||||||||||||||||||||||||| ||
Db   4260 TTGTTTTTTACCACTTCTAACTAAATTTAATCATTTTATTCCCAATTGCGAATAGAGCCATAA 4319

Qy   4320 accaaagtgggtaataagggtacctgtgattttgtcattcatcccaatacacagacat 4379
          |||||||||||||| ||||||| |||| |||||||||||||||||||||||||||||
Db   4320 ACTCAAAGTGGGTAATAAAGGTACCTGTGATTTTGTCATTACCAATAGAAATCACAGACAT 4379

Qy   4380 tttatactatatcatagttgttcaggtaacgttcgtatgtgaaatatattctatccaaaact 4439
          ||||||||||| ||||||||||||| |||||||||||||||||||||||||||||||| ||
Db   4380 TTTATACTATATTACAGTTGTTGCAGTACGTTGTTGAAGTATATTTATACTCAAAACT 4439

Qy   4440 acttgaaattagaacctcctgctggaatccgtgctgtttttaacatattaaataaaacatgttttaa 4499
          |||||||||||||||||||||| ||||||||||||||| ||||||||||||||| |||||
Db   4440 ACTTTGAAATTAGAACCTCCTGCGGGATCTTGTTTTTAACATATTAATAAAAACATGTTTAA 4499
```

```
Qy   4500 aattttgatattttgatataatcatatttcattcattgtttcctttgtttgtaatctatattt 4559
          |||||||||| || ||||||||||||||||||||||||| |||||||||||||||| ||||||
Db   4500 AATTTTGATATTTTTGATATAATCATATTTCATTATCATTTGTTTCCTTTGTAATCTATATTT 4559

Qy   4560 tatatatttgaaaacatcttcttctgagagagttcccaagagttcaccaatgagggttccttg 4619
          |||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||||
Db   4560 TATATATTTGAAAAACATCTTTCTTCTGAGAAGAGTTCCCAGAGTTCACCAAATGAGGGTTCCTTG 4619

Qy   4620 gcatgcacacacacagagtaagactgatttctaggacattgtacctgctgag 4679
          ||||||||||||||||||||||||||||||| || |||||||||||||||||
Db   4620 GCATGCACACACACAGAGTAAGAACTGATTTTAGAGGCTAACAATTGCACATTGTGCCTGAG 4679

Qy   4680 atgcaagactgaaattagaaaaagtttctccaaagtacacagtacacagtggctgacctg 4739
          ||||||| |||||| ||||||||||| |||||||||||||||||||||||||||||||||
Db   4680 ATGCAAGACTGAAATTAGAAGAAGTCTCCAAAGATACACAGTACACAGTGGTGCCTGAG 4739

Qy   4740 gaggggggaaatctccgcgcttctacctag gtaatgcctcctatagggagggctgggcttcgt 4799
          |||||||||| |||||||| ||||||||||||||||||||||||||| ||||| ||||||
Db   4740 GAGGGGGGAAAATCTCCGCGCTTCTACAATGCGTCCCCGGGAGGGCTGGGCTGCTGT 4799

Qy   4800 ccttgtgttctggctgtgctgttatttttttctgtctccctgctccctacgtcttcaaggacttgtt 4859
          |||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db   4800 CCTTGTGTTCTGGCTGTGCTGTTATTTTTTCTCTGTCTCCTCGTTTAAAGGACTTGTT 4859

Qy   4860 tggatctccagttccgttcctcctagcatagtgcctggacaagtgcaggttcttcaaatgagtttgcaga 4919
          |||||||||||||| |||||||||||| ||||| |||||||||||||||||||||||||
Db   4860 TGGATCTCCAGTTCCGTTCCTAGCATAGTGCCTGGCACAGTGCAGGTTCTCAAATGAGTTTGCAGA 4919

Qy   4920 gtgaatgggaaatataaaactagaaatatatatctgttgaaatcagacaccaccagtagtcctg 4979
          ||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db   4920 GTGAATGGGAAATATAAAAAATTAGGGTGCATAAAAATATCCTTGGTTTGAAATCAGACACCACCAGTAGTCCTG 4979

Qy   4980 gtgtaagtgtgtacgtgtcgtgtgtgtgtgtgtgtgtgtcaacaggtgg 5039
          |||||||||||| |||||| ||||||||||||||||||||||||||||||
Db   4980 GTGTAAGTGTGTAC----GTGTGTGTGTGTGTGTGTGTGTGTAAAACAGGTGG 5035

Qy   5040 agaataggagaactattattgggggtatgggtgcataaattgggaatggggatggttcttttttaaaga 5099
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5036 AGATATAGGAGAACTATTATTGGGGGTATGGGTGCATAAAATTGGGATGGTTCTTTTTTAAAGA 5095

Qy   5100 aactccaaaacagacttctggaaggttattttttcattttctctaagaaatcttgctggcagcgtgaagcca 5159
          ||||||||||||||||| |||||||||||||||||||| |||||||||||||||||||
Db   5096 AACTCCAAACAGACTTCTGGAAGGTTATTTTCTAAGAATCTTGCTGGGCAGCGGGCA 5155

Qy   5160 acccccctgtgcacagcgccacccagcgccaccacgtctcctcccccatgaag 5219
          |||||||| |||||||||| ||||||| |||||||| || |||||| |||||||
Db   5156 ACCCCCTGTCCACACGTGCCCACCCTCGCCACTCTGTCTTCCCCCATGAAG 5215

Qy   5220 ggctggctcccagtatatataaacctctctgggagctccgggcacgtgagcgcaggcaaggccac 5279
          |||||||||||||| || |||| |||||||||||||||||||||||||||||||
Db   5216 GGCTGGCTCCCCAGTATATATAAACCTCTCTGGGAGCTCCGGGCATGAGCGCAGGCCAC 5275

Qy   5280 ccatccaaggcaccctcctcagcacac 5304
          |||||||||||||||||||||||||||
Db   5276 CCATCCAGGCACCCTCCTCAGCACAGC 5300
```

```
REFERENCE   2  (bases 1 to 170425)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkly,L.,
            Bukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Illev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J.C., Landers,T., Largocque,R., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Scheuer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassilliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            -------------------------------------------------------
COMMENT     On May 25, 2000 this sequence version replaced gi:7249345.
            All repeats were identified using RepeatMasker:
            Smit,A.F.A. & Green,P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------------------------------------------------
            --------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            --------------- Project Information
            Center project name: L7153
            Center clone name: 138_F_3
            --------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 151483 bases at least Q40
            Consensus quality: 160439 bases at least Q30
            Consensus quality: 164708 bases at least Q20
            Insert size: 167000; agarose-fp
            Insert size: 167725; sum-of-contigs
            Quality coverage: 4.1 in Q20 bases; agarose-fp
            Quality coverage: 4.1 in Q20 bases; sum-of-contigs
            ---------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 28 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1 1300: contig of 1300 bp in length
            *     1301 1400: gap of 100 bp
            *     1401 2412: contig of 1012 bp in length
            *     2413 2512: gap of 100 bp
            *     2513 4147: contig of 1635 bp in length
            *     4148 4247: gap of 100 bp
            *     4248 5455: contig of 1208 bp in length
            *     5456 5555: gap of 100 bp
            *     5556 6997: contig of 1442 bp in length
            *     6998 7097: gap of 100 bp
            *     7098 8534: contig of 1437 bp in length
            *     8535 8634: gap of 100 bp
            *     8635 10164: contig of 1530 bp in length
            *    10165 10264: gap of 100 bp
            *    10265 10487: contig of 223 bp in length
            *    10488 10587: gap of 100 bp
            *    10588 12420: contig of 1833 bp in length
            *    12421 12520: gap of 100 bp
            *    12521 14843: contig of 2323 bp in length
            *    14844 14943: gap of 100 bp
            *    14944 17265: contig of 2322 bp in length
            *    17266 17365: gap of 100 bp
            *    17366 20071: contig of 2706 bp in length
            *    20072 20171: gap of 100 bp
            *    20172 23531: contig of 3360 bp in length
            *    23532 23631: gap of 100 bp
            *    23632 26137: contig of 2506 bp in length
            *    26138 26237: gap of 100 bp
            *    26238 29858: contig of 3621 bp in length
            *    29859 29958: gap of 100 bp
            *    29959 36225: contig of 6267 bp in length
            *    36226 36325: gap of 100 bp
            *    36326 44015: contig of 7690 bp in length
            *    44016 44115: gap of 100 bp
            *    44116 49493: contig of 5378 bp in length
            *    49494 49593: gap of 100 bp
            *    49594 56796: contig of 7203 bp in length
            *    56797 56896: gap of 100 bp
            *    56897 62860: contig of 5964 bp in length
            *    62861 62960: gap of 100 bp
            *    62961 72552: contig of 9592 bp in length
            *    72553 72652: gap of 100 bp
            *    72653 83310: contig of 10658 bp in length
            *    83311 83410: gap of 100 bp
            *    83411 92992: contig of 9582 bp in length
            *    92993 93092: gap of 100 bp
            *    93093 103886: contig of 10794 bp in length
            *   103887 103986: gap of 100 bp
            *   103987 119418: contig of 15432 bp in length
            *   119419 119518: gap of 100 bp
            *   119519 136695: contig of 17177 bp in length
            *   136696 136795: gap of 100 bp
            *   136796 153977: contig of 17182 bp in length
            *   153978 154077: gap of 100 bp
            *   154078 170425: contig of 16348 bp in length.
FEATURES             Location/Qualifiers
     source          1..170425
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1"
                     /clone="RP11-138F3"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1300
                     /note="assembly_fragment"
     misc_feature    1401..2412
                     /note="assembly_fragment"
     misc_feature    2513..4147
                     /note="assembly_fragment"
     misc_feature    10265..10487
                     /note="assembly_fragment"
     misc_feature    8635..10164
                     /note="assembly_fragment"
     misc_feature    7098..8534
                     /note="assembly_fragment"
     misc_feature    5556..6997
                     /note="assembly_fragment"
     misc_feature    4248..5455
                     /note="assembly_fragment"
                     /clone_end:SP6
                     /vector_side:right"
     misc_feature    10588..12420
                     /note="assembly_fragment"
     misc_feature    12521..14843
                     /note="assembly_fragment"
     misc_feature    14944..17265
                     /note="assembly_fragment"
     misc_feature    17366..20071
                     /note="assembly_fragment"
```

```
misc_feature   20172..23531
               /note="assembly_fragment"
misc_feature   23632..26137
               /note="assembly_fragment"
misc_feature   26238..29858
               /note="assembly_fragment
               vector_side:left"
               clone_end:T7"
misc_feature   28959..36225
               /note="assembly_fragment"
misc_feature   36326..44015
               /note="assembly_fragment"
misc_feature   44116..49493
               /note="assembly_fragment"
misc_feature   49594..56796
               /note="assembly_fragment"
misc_feature   56897..62860
               /note="assembly_fragment"
misc_feature   62961..72552
               /note="assembly_fragment"
misc_feature   7653..83310
               /note="assembly_fragment"
misc_feature   83411..92992
               /note="assembly_fragment"
misc_feature   93093..103806
               /note="assembly_fragment"
misc_feature   103987..119418
               /note="assembly_fragment"
misc_feature   119519..136695
               /note="assembly_fragment"
misc_feature   136796..15397
               /note="assembly_fragment"
misc_feature   15078..170425
               /note="assembly_fragment"

BASE COUNT     48243 a    34101 c    34753 g    50621 t    2707 others
ORIGIN

Query Match  98.5%;  Score 5223;  DB 77;  Length 170425;
Best Local Similarity 99.5%;  Pred. No. 0;
Matches 5283;  Conservative 0;  Mismatches 15;  Indels 11;  Gaps 4;

Qy      1 atctttgttcagttta--cctcagggtcttattatgatgaaatgaagataaccaatgttgaa 58
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db  63685 ATCTTTGTTCAGTTTATCCTTACAGGGCTATTATGAAATGAAGATAACCAATGTGAA 63744

Qy     59 agtcctataaactgtatagcctccattcg-gatctatgtgtcttttggcagg-atgataaag 115
          ||||||||||||||||||||| ||||| |||||||||||||||||||||||| ||||||||||
Db  63745 AGTCCTATATACTGTATTACCCTCCATTCGTGATGTATGTCTTTGGCAGGTATGTATATAAAG 63804

Qy    116 aatcaggaggaaggagtatccacgttagccaagtgtccaggctgtgtctgctctcttattt 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  63805 AATCAGGAGGAAGGAGTATCCACGTTAGCCAAGTGTCCAGCCTGTGTCTGTGCTCTTATTTT 63864

Qy    176 agtggacagatgtgtcctgacagaagctattcttcaggaaacatcaactccaatatggt 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  63865 AGTGACAGAGTTGCTCCTGACAGAAGCTATTCTTCAGGAAACATCAACATCCAATATGGT 63924

Qy    236 aaatccatccaaacaggagccaaacagaaatcagacatcgggaagaaactaaaatcaggaaa 295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  63925 AAATCCATCCAAACAGGAGCTAAGAAACACAGGAAATCAGACATGGGCACTTGCCCAAGGAAAAAT 63984

Qy    296 gccaggagggagcaaataatgatgaaaataaacttttccctttgttcttttaattttccagga 355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  63985 GCCAGGAGGGAGCAAATAATGATGAAAATAAACTTTTCCCTTGTTCTTTTAATTTCAGGAAG 64044

Qy    356 aaatgatgaggaccaaaatcaaatgaagaatgaggaatatacttcaatgacaaaaagat 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64045 AAATGATGAGGACCAAAATCAAATGAAGAATGAGGAATATACTTCAATGACAAAAAGATTTCAA 64104

Qy    416 atggtaattaagtattttgttccttgggaagagagagagacctccatgtgagcttgatgggaaaat 475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64105 ATTGGTAATTAAGTATTTTGTTCCTTGGGAAGAGAGAGACTTCCATGTGAGCTTGATGGGAAAAT 64164
```

```
Qy    476 gggaaaacgtcaaaagcatgatctgatcagatccaaagtggatttatattaagac 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64165 GGGAAAAACGTCAAAAGCATGAATCGATCTGATCAGATCCCAAAGTGGATTATATTTAAAAAC 64224

Qy    536 cagatggcatcactctgggggaggcaagttcaaggaaagtcatgttagcaagcaacaaca 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64225 CAGAATGGCATCACTCTGGGGAGGCAAGTTCAGGAGAGGTCAITGTTAGCAAAGGACAATAACA 64284

Qy    596 ataacagcaaaatcaaaattcccgcaaaatgcaggaggagaaaaatgggggaaagcttc 655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64285 ATAACAGCAAAATCAAAAATTCCGCAAAATGCAGGAGGAGAAAAATGGGGGAAAGCTTTC 64344

Qy    656 ataacagttgattaggcagttgcaacacatgtttcgccaacaacctccccgtctataccaagggaaca 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64345 ATAACAGTGATTAGGCAGTTGACCAITGTTCGCAACACCTCCCGTCTAAACCAGGGAAAA 64404

Qy    716 caaaaattgactgggctaagcctggacttcaagggaaatatgaaagagcaaaa 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64405 CAAAAAATTGACTGGGCTAAGCCTGGACTTTCAAGGGAAATATGAAAAAACTGAGAACAAAA 64464

Qy    776 caaaagacatgggttaaaaaggcaaccaagaacattgttgagacctgagcttgaagctt 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64465 CAAAAGACAITGGTTAAAAAGGCAACCAAGAACAITGTGAGACCTTCAAGGCCTTCAAAGGCCCT 64524

Qy    836 cagcaggggacctgagggcaitttgcctttaggaaagccagttttcttaaggaaatcttaaga 895
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64525 CAGCAGGGGACCCTGAGGGCAITTTGCCTTTAGGAAAGCCCAGTTTTCTTAAGGAAATCTTAAGA 64584

Qy    896 aactcttgaaagtacatgaatttaaccattttaagtataaaacaaaatatgcgatgcata 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64585 AACTCTTGAAAGATCATGAAITTTTAAGGTATAAAACAAAATATGCGATGCATA 64644

Qy    956 atcagtttagacatgggtcccaaattttataaaagtcaaggcatacaaggataacgtgtccca 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64645 ATCAGTTTAGACATGGGTCCCAAATTTTTATAAAGTCAAGGCAITACACAAGGATAAACGTGTCCCA 64704

Qy   1016 gctccggataggtcagaaatcattagaaatcactgtgtcccatcctaacttcttccagaa 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64705 GCTCCGGGATAGGTCAGAAAITCAITAGAAAITCACTGTGTCCCATCCTAACTTTTTCAGAA 64764

Qy   1076 tgatctgtcatagccccgatcgtgtctgacctccaacacaacctacaa 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  64765 TGATCTGTCATAGCCCCGATCGTGTCTGACCTCCAACACACAACTACAA 64824

Qy   1136 cccaagtgcctcaaccaitgttaaggtgtcaatccaagtaggtccaitgccacc 1195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64825 CCCAAGTGCCTCAACCAITGTTAACGTGTCAITCCCAITTACAACACCATCTACAG 64884

Qy   1196 tcccctgtgcagccaitcccgcctccaaggaagtctccccactctagacttctgcatccac 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64885 TCCCCTGTGCAGCCCAITCCCGCCTCCAAGGAAGTCTCCCCACTTAGACTTCTGCATCAC 64944

Qy   1256 gatgttacagccagaagctccgtgaggtgaggtgtcttgtcttacacctaccgtatgc 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64945 GATGTTACACGCCAGAAGCTCCGTGAGGTGTGTCTTACACCTACCGTATGC 65004

Qy   1316 tctacacctgagctcaacctctgcctccaaggtcaagcaaitcccctgctcag 1375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65005 TCTACAACCTGAGCTCAACTGCCTCCAAGGTTCAAGCAAITTCTCCCGTTCTCAG 65064

Qy   1376 cctcccgcgtagctggggactacaggcgcacgcccggctaaitttgtattgttagtagag 1435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65065 CCTCCCGCGTAGCTGGGACTACAGGCGCCACGCCCGGCTAAITTTTGTAITTGTTAGTAGAG 65124

Qy   1436 atggggtttcaccatattagcccggctgtcttgaactcctgacctcaggtgatccacc 1495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65125 ATGGGGTTTCACCATAITTAGCCCGGCTGGTCTTGAACTCCTGACCTCAGGTGAITCCACC 65184

Qy   1496 acctcagcccctaaagtgctggattacaggcatgagtcaccgcccggccaagggtc 1555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  65185 ACCTCAGCCTCCTAAAGTGCTGGGATTACAGGCATGAGTCACCGCCCGGCCAAGGGTC 65244
```

```
Qy    1556 agtgtttaataaggaataacttgaaatgttactaaaccacaagggaaacgacaaagc 1615
           ||||||||||||||||||||||||||| || |||||||||| |||||||| |||||||
Db   65245 AGTGTTTAATAAGGAATAACTTGAATGGTTTACTAAACCAACAGGGAAACAGACAAAGC 65304

Qy    1616 tgtgataatttcaggaattcttgggatggggaatggtgccatagagctgcctagtcc 1675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65305 TGTGATAATTTCAGGAATTCTTGGGATGGGGAATGGTGCCATAGAGCTGCCTAGTCC 65364

Qy    1676 cagaccactggtcctcatcacttcttcccctccatcccaccttccaggctagttaccatt 1735
           ||||||||||||||||||||||||||||||||||| |||   ||||||| ||||||||||
Db   65365 CAGACCACTGGTCCTCATCACTTCTTCCCCTCCATCCCATTTTCAGGGTAAGTTACCATT 65424

Qy    1736 ttattcaccatgcttttgtggttaagcctccaacatcgttactgaaaatgaactataa 1795
           ||||||||||||||||||||||||||||||||||||||| |||||||||  ||| ||||
Db   65425 TTATTCACCATGCTTTTGTGGTTAAGCCTCCAACATCGTTACTGAAAATAAGAGTATACATAA 65484

Qy    1796 actagttccatttgggggccatctgttgtgttgttgttatagggggcatacccagaga 1855
           || ||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db   65485 ACTAGTTCCATTTGGGGCCCATCTGTTGTGTGTTATAGGGGGCATACCCAGAGA 65544

Qy    1856 ctccttgaaggccccccggccagaggtttccttcctcctccaagccaccccg 1915
           ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   65545 CTCCTTGAAGGCCCCCGGCCAGAGGTTTCCTTCCTCCTCCCAAGCCACCCGG 65604

Qy    1916 ggtcctgggtgtcctgagcaactgacctgccaactcgggggagcccctgccaactct 1975
           ||||||||||||||||||||||||| ||||||| || |||||||||||| ||||||
Db   65605 GGTCCTGGGTGTCCTGACAACTGACAGTTTTCCTGCCAACTGGGGGAGCCCTGCCAAGTGTTTATCACTTCT 65664

Qy    1976 ctaggggacctgttgcttctattttctgtgttgactcgttcattcaatccaggtcattcatga 2035
           ||||||||||||||| |||||||| ||||| |||||| |||||||||||||| |||||||||
Db   65665 CTAGGGACCTGTTGCTTTCTATTTCTGTGTGTGACTCGTTCATTCATTCAGGCATTCATTGA 65724

Qy    2036 caatttattgagtactttatatctgccagacacccagagagacaaaatggttggca 2095
           ||||||||||||||||||||||||| |||||| |||||||||||| |||||||||
Db   65725 CAATTTATTGAGTACTTTATATCTGCCAGACACCAGAGAGACAAAAATGGTTGGCA 65784

Qy    2096 actgcctacctctcgtgggtgacagttcatgtgaagagactgcagaagaaaataat 2155
           |||||||||||||||||||||||||| ||||||||||||||| |||||| ||||||
Db   65785 ACTGCCTACCTTCGTGGAGTGACAGTTTCATGTGAAGAGACGTGCAGAGAAAAATTTAAT 65844

Qy    2156 agccagccaacttaaacccagtgctgaagaagaaggaaataacacccatctctgaagatg 2215
           |||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||
Db   65845 AGCCAGCCAACTTAAACCCAGTGCTGAAAGAGGAAATAAACACCACTCTTGAAAGATTG 65904

Qy    2216 tgcgcagcatcccttacaagggcaacctcctaacgcgccccctgcctgccaacctgtgcc 2275
           |||||||||||||||||||||||||||| |||| ||||| |||||||||||||||||||
Db   65905 TGCGCAGGATCCCTTAACAAAGGCCACCTCCCACGCGCCCCTGCCTGCCTCCATCGTGCCC 65964

Qy    2276 ggagggccccccaagcacccaagctctctcccaatcagtcacagcctgtcagct 2335
           || ||||||||| |||||| ||||||||| |||| ||||| ||||||| |||||
Db   65965 GGAGGGCCCCCAGTCTTCCAACCCGAGTCTTCCAACCTCAGTCACGCGCTGCAGCT 66024

Qy    2336 ggcctgcctcgcttcccgtgatcgtctgttgcatctgagctggagactcccttggctcc 2395
           ||||||||||| ||||||||||||||||| ||| ||||||||||||| ||||||||||
Db   66025 GGCCTGCCTCGCTTCCCGTGAATCGTCTCCTGGTCATCGAGCTGGAGACTCCTTGGCTCC 66084

Qy    2396 aggcccagaagggaaatgagggagaagctagtctgaggggacagt 2455
           ||||||||||||||| |||| |||||||| ||| |||||||||||
Db   66085 AGGCTCCAGAAAGGGAAATGGAGGGGAAATTGACGAGGAATCTAACGGAGAATCTGGAGGGGACAGT 66144

Qy    2456 gtttccctccagagggaaagggcctcccacgtccaggaggaggtggggactgca 2515
           ||||||||||| |||| |||| |||||||||||||||||||||||| ||||||
Db   66145 GTTTCCTTCGAGGGGAAAGGGCCTCCACGAATTCCAGGAGGTGGGGACTGCA 66204

Qy    2516 gggagtgggagctgggagctgaaggggcaagg 2575
           |||||||||||| |||||||||||||||||||
Db   66205 GGGAGTGGGAGCTGCTGAAGCGCTGAAGGGCTGAAAAGGGCAAGG 66264

Qy    2576 ctgaagctgcccagagtgttcagcggggctgggagtttccgttcctcctg 2635
           ||||||||| ||||| |||| ||||||||| |||||| |||||||||||
Db   66265 CTGAAGTGCCCAGATGTTCAGTGTTGTTCACGGGACTTTTTCGTTGCTTCCTG 66324

Qy    2636 tgagcccttttatcttctcgcttgggagagaagaagtcattcatgaaggggatgca 2695
           |||||||||| || |||| | |||||||||||||||||||| |||||||| ||||||
Db   66325 TGAGCCTTTTTATCTTGGAGGAGAAGAAGTCTATTTCATGAAGGGATGCA 66384


Db   66325 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TGAGCCTTTTTATCTTGGAGGAGAAGAAGTCTATTTCATGAAGGGATGCA 66384

Qy    2696 gtttcaataaagtcagctgttcaaaattccagggtgtgtcagtggtttcttcctgaaggc 2755
           ||||| ||||||||||| ||| ||||||||||| ||||| ||| |||||||||||| |||
Db   66385 GTTTCATAAAAGTCAGCTGTTAAAATTCCAGGGTGTGCAGTTGTTTCTTCCTTCACGGAAGGCC 66444

Qy    2756 tttattttaatgggaatatataggaagcgagctcattccattccagatcacggaagaa 2815
           || ||||| | || |||||| ||||| || |||||||||||||||| || || |||||
Db   66445 TTTATTTTAATGGGAATATATAGGAAGCGAGCTCATTCCATTCCAGGTCACGGAAGAA 66504

Qy    2816 gtggactggagtctcacatgtctctgactgtgccaagtgtcggttaattcacggagca 2875
           ||||||||||| |||||||||||||| |||| |||||||| | || |||||||||||
Db   66505 GTGACTGGGAGTCTCACGTCTTTCTTCATGTCTCAGCCCGTGGCAACCCTGTGGGTGGACTTG 66564

Qy    2876 gcttatgtcaaagacgggtcgaaaacccttgaatccagggtgtttttcttgttctg 2935
           ||||| ||||||| |||||| || || ||||||| |||||||||||| ||||||||
Db   66565 GCTTATGTCAAGACGGTCGAAAACCTTGGAATTCAGGAGACTCGGTTTTCTTTCTGTTCTG 66624

Qy    2936 ccattggttggctgtgtcgaccgttgggcaagtgttctctcatcagtctcctcactgatctcagtca 2995
           |||||| |||||| ||| | ||| |||||||| |||| || || ||| || ||| || |||
Db   66625 CCATTTGTTGGCTGTGTCGACCGTGGGCAAGTGTCTCTCTTCACTGATCACGTCA 66684

Qy    2996 ctgctataaagaccctttgcagctctcggttgttctctgtgaacacttcctgtg 3055
           |||||||||||||||||| ||||||| |||| ||| |||||||||||| |||||
Db   66685 CTGCTATAAAGACCCTTGCAGCTCTCGTGTTCTGTGAACACTTCCCTGATTCTCTGTG 66744

Qy    3056 aggggggatgttgagggaggggaggcagagggaggggacacagggggtg 3115
           ||||||||| || |||||||||||||| ||||||||||||| ||||||
Db   66745 AGGGGGGATGTTGAGGGAGGGGAGGCAGAGGGGAAGGCAGAGGGGACGTGGAGGTG 66804

Qy    3116 gaggggggacagaggaaggcagcagagaagctgggctggtggtcttccatggccagga 3175
           |||||||||| |||||||||||||| |||||| |||||||| |||||| ||||||||
Db   66805 GAGGGGGACAGAGGAAGGCAGCAGGGAAGCTGGGTGGTGGTCTTCCATCAGTCCAGA 66864

Qy    3176 gactccaggacccgagcagcacaatgcttcaggaaagctcaatgaacccaacccacatt 3235
           || ||||||||||||||||| |||| ||| || ||||| |||||||||||| |||||||
Db   66865 GACTCCAGGACCGAGCAGCACCAATGCTTCAGGAAAGCTCAATGAACCAACAGCCACATT 66924

Qy    3236 ttccttcccctagcatagcaatgggcgcgggggaaaagagactga 3295
           || ||| || ||| |||| ||| |||| |||| ||||||||||
Db   66925 TTCCTTCCCTAAGCATAGACAAATGGCAATAACCAAACAAAGAATGGCAGAGGTGGGAGGACTAA 66984

Qy    3296 ctgtggttagtcttctgccctggctcattcaaaaaactgggccaaggacaagtggaaaatgccaga 3355
           ||| ||||| |||||||||||||| ||||||||||||||||| |||||| ||||||||||||
Db   66985 CTGTGGTTAGTCTTTTGCCCTGGCATTCAAAAAACTGGGCATTCAGAGCAAGAAGTGAAAAATGCCAGA 67044

Qy    3356 gattgttaaacttttcacccctgaccagcaagtgactgctgacag 3415
           ||||| |||||||| ||||||| ||||||||||||| |||||||
Db   67045 GATTGTTAAAACTTTCACCCCTGACCAGCAAGTGACTCGCTGACAG 67104

Qy    3416 cacggagtgacccagcgacgcgaggaggagaaggagaagagtagtgtatgagcaa 3475
           || |||||||| || ||| || ||| ||||||||||||||||||||| ||| |||
Db   67105 CACGGAGTGACCTGCAGCCGCGGAGGAGAAGGAGAGAGGAGATAGTGTATGGAGCAA 67164

Qy    3476 gaaagaacaagatcattcaagcatgtggaattgacccacagaggattatagtccacgtgat 3535
           ||||||||| || || || ||| || |||| ||| |||||||||||| ||||||| ||||
Db   67165 GAAAGACAGATTCATTCAAGGCCAGTGGGAATTGACCCACAGAGGATTATAGTCCACGTGAT 67224

Qy    3536 cctgggttcctaggagggcaggggctatattgtgggggaagtcg 3595
           |||||||||| |||||| |||| |||||| ||||||||||||||
Db   67225 CCTGGGTTCTAGGGAGGGCAGGGCTATATTGTGGGGGAAGTGGGGGAAAATCAGTTCAAGGGAAGTCG 67284

Qy    3596 gggaacctgattctaatactatattttcccttaacagctggtgataactctggagtc 3655
           ||||| |||||| ||| |||| || |||||| |||||||| |||||| ||||||||
Db   67285 GGAGACCTGATTTCTAATACTATATTTTCCTTTACACTGAGTAAATTCGAGACAAGTC 67344

Qy    3656 acaaggtagtaactgaaggctgtaagattacttagttctcccttattcttttc 3715
           |||||||| |||||| | ||| ||||| || ||||| |||| ||||| |||||
Db   67345 ACAAGGTAGTAACTGAAGGCTGTAAGATTACTTAGTTTCTCCTTATTTAGGAGAACTCTTTTTC 67404

Qy    3716 tctgtgggagttagcacaaggcaatcccgtttcttttaacaggagaaaacattcct 3775
           ||||||||| || |||||||||||| || |||||| ||||||||||||| |||||||
Db   67405 TCTGTGGGAGTTAGCACAAGGCAATCCCGTTTCTTTTAACAGGAGAAAAACATTCCT 3775
```

```
Db   67405 TCTGTGGAGTTAGCAGCACAAGGGCAATCCCGTTTCTTTAACAGGAGGAGAAAAACATTCCT 67464

Qy    3776 aagagtaaagcccaaacagagattccaaggctaggtcttgctgactatatgatggtttttga 3835
           |||||||||||||||||| |||||| |||||||||||||||||||||||||||||||||||
Db   67465 AAGAGTAAAGCCCAAACAGAGATTCAAGGCTAGGTCTTGCTGACTATATGATGGTTTTTGA 67524

Qy    3836 aaaatcaattcaaggcgatgttacatctgatctgattcagagaaatgagactagtaccctttggtc 3895
Db   67525 AAAATCAATTTCAGCGCGATGTTTACTATCTGATTCAGAGAAATGAGACTAGTACCCTTTGGTC 67584

Qy    3896 agctgtaaacaacaacccccatttgtaaatgtcctcaagttcaggcttaactgcagaaccaat 3955
Db   67585 AGCTGTAAACAACAACACCCATTTGTAAATGTCCTCAAGTTCAGGCTTAACTGCAGAACCAAT 67644

Qy    3956 caaataagaaatagaatctcttagagcaaactgtgttcctccactctggaggtgagtcgtgcc 4015
Db   67645 CAAATAAGAAATAGAATCTTTAGAGCAAACTGTGTTCTCCACTCTGGAGGTGAGTCGTGCC 67704

Qy    4016 agggcagtttggaaatattttactcacaacatgtattgttctccactctgtgttaacaacat 4075
Db   67705 AGGGCAGTTTGGAAATATTTTCACAACATGTATTGTTCTCCACTCTGTGTTAACAACAT 67764

Qy    4076 aaagttgctccaaaggcaatcattattttactccaagttactctctgacacagttttgg 4135
Db   67765 AAAGTTGCTCCAAAGGCAATCATTATTTTACTCCAAGTTACTCTCTGACACAGTTTTGG 67824

Qy    4136 tatatttatggctattcgccattcgcttttcttttgggttttattttattaatgtaa 4195
Db   67825 TATATTTATGGCTATTCGCCATTCGCTTTTCTTTTGGGTTTTATTTTATTAATGTAA 67884

Qy    4196 agcagggatattattaacctacagtccagaaaagccctgtgtgaatttgaatgaggaaaaattac 4255
Db   67885 AGCAGGGATTATTAACCTACAGTCCAGAAAAGCCTGTGAATTTGAATGAGGAAAAATTAC 67944

Qy    4256 gttttattttacccctcctaactaaattaacattttattccattgccgaatagagagcc 4315
Db   67945 ATTTTGTTTTTACCACCTTCTAACTAAATTTATTCCATTGCCGAATAGAGCC 68004

Qy    4316 ataaaactcaagtggtaataagagtacctgtgatttgtcattaccaatagaaatacacg 4375
Db   68005 ATAAACTCAAAGTGGTAATAACAGTACCTGATTTTGTCATTACCAATAGAAATCACAG 68064

Qy    4376 acaatttatactatattacaagtcgttgttgcaggtaagtgaaaatattatactcaa 4435
Db   68065 ACATTTATATACTATATTACAGTTGTTGCAGGTAACGTGTAAGTGAAATATTATACTCAA 68124

Qy    4436 aactactcttgaaattagactccctcggatctgtgttcattaatatattaataaacacagt 4495
Db   68125 AACTACTTTGAAATTAGACTCCCTGGATCTGTGTTCATTAATATATTAATAAACACATGT 68184

Qy    4496 ttaaaaatttgatatttgataatacatatttcattcattctgtttcttttgtaatctat 4555
Db   68185 TTAAAATTTGATATTTGATAATCATATTTCATTCATTCTGTTTCCTTTGTAATCTAT 68244

Qy    4556 atttatatatatttgaaaaaatcttctctgagaaggaagagttccccaatgagagtt 4615
Db   68245 ATTTTATATATTTGAAAACATCTTTCTCTGAGAAGAGTTCCCAATGAGAGTT 68304

Qy    4616 cttggcatgcacacacagtaagaacctgattggaggctaacactggtggtgcc 4675
Db   68305 CCTTGGCATGCACACACAGTAAGAACTGATTTAGAGGCTAACACATTGACATTGCGCC 68364

Qy    4676 tgagtgcaagactagaattagaaagttctccccaaagatacacagtgttgtcttaaagctag 4735
Db   68365 TGAGATGCAAGACTGAAATTAGAAAGTTCTCCCAAAGATACACAGTTGTTTTAAAGCTAG 68424

Qy    4736 gggtgagggggaaatctgccgcctctataggagatacacagttgtttgttgagggttg 4795
Db   68425 GGGTGAGGGGGAAATCTGCCGCGTTCTATAGGAATGCTCTCCCTGGACGCCTGGTGAGGGTG 68484

Qy    4796 ctgtcccttgtgttctggctgtcgttattttcttcctgctccctgctacgtcttaaaggact 4855
Db   68485 CTGTCCTTGTGTTCTGGCTGTCGTTATTTTCTTCCTGCTCCCTGCTACGTCTTAAAGGACT 68544

Qy    4856 tgtttggatctccagtcctagcatatagtgcctggcacagtgcacagtgttctcaatgtttg 4915
Db   68545 TGTTTGGATCTCCAGTTCCCAGTTCCTCAGCATAGTAGTGCCTGGCACAGTGCACAGTGTTG 68604

Qy    4916 cagagtgaatggaaatataaactagaaatatatatctcttgttgaaatcaagcacacagtagt 4975
Db   68605 CAGAGTGAATGGAAATATAAAACTAGAAATATAAATCCTTGTTGAAATCAAGCACACCAGTAGT 68664

Qy    4976 cctgtgttaagtagtgtgtgtgtgtgtgtgtgtgtaaaaccag 5035
Db   68665 CCTGGGTGTAAGTGTGTGTAC------GTGTGTGTGTGTGTGTGTTAAAACCAG 68718

Qy    5036 gtgggagatataggaactacattatttgggtatggtgcaataaattgggatgttcttttttaa 5095
Db   68719 GTGGGAGATATAGGAACTACTATTTATTTGGGGTCATAAAAATTGGGATGTTCTTTTTAA 68778

Qy    5096 aagaaactccaaacagagactcctggaagttcacgtgtcttgggaaggttattttcaacat 5155
Db   68779 AAGAAACTCCAAACAGAGACTTCCGGAAGGTTATTTTCTTAAGAATCTTGCTGGCAACAT 68838

Qy    5156 ggcacccccccctgtgtcacacccccccccacccagcctcacgtggccacctctgtcttcccccat 5215
Db   68839 GGCAACCCCCCTGTCGACACAGCCCCCACCCAGCCTCACGTGGCCACCTCTGTCTTCCCCAT 68898

Qy    5216 gaagggctggcgtctcccagtatatatataaacctctctggagtctcggggcaatgagccagcaagg 5275
Db   68899 GAAGGGCTGGCGTCTCCCAGTATATATATAAACCTCTCTGGAGCTCTCGGGGCATGAGCCAGCAAGG 68958

Qy    5276 ccacccatcccggccacctcctcagcacacagc 5304
Db   68959 CCACCCATCCAGGCACCTCTCAGCACAGC 68987
```

RESULT 4
HSMYOC1

```
LOCUS       HSMYOC1    2800 bp    DNA    PRI    21-APR-1998
DEFINITION  Homo sapiens myocilin (GLC1A) gene, promoter and exon 1.
ACCESSION   AF049791
VERSION     AF049791.1  GI:3065670
KEYWORDS    .
SEGMENT     1 of 3
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2800)
  AUTHORS   Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C.,
            Alward,W.I.M., Sheffield,V.C. and Stone,E.M.
  TITLE     Characterization and comparison of the human and mouse GLC1A
            glaucoma genes
  JOURNAL   Genome Res. (1998) In press
REFERENCE   2  (bases 1 to 2800)
  AUTHORS   Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C.,
            Alward,W.I.M., Sheffield,V.C. and Stone,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-1998) Ophthalmology, University of Iowa, 200
            Hawkins Drive, Iowa City, IA 52242, USA
FEATURES             Location/Qualifiers
     source          1..2800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="I"
                     /map="1q21-q31"
     promoter        1..1905
                     /gene="GLC1A"
     exon            1906..2509
                     /gene="GLC1A"
                     /number=1
BASE COUNT      781 a    588 c    673 g    758 t
ORIGIN
```

Query Match

Best Local Similarity   34.6%;   Score 1836;   DB 66;   Length 2800;

Matches 1864;   Conservative   99.5%;   Pred. No. 0;   Mismatches 0;   Indels 5;   Gaps 2;

```
Qy  3431 agcgcgcaggggaggaggaagaagaaagaagagggagataagtcgtatgacaagacaagattcat 3490
         ||||||||||||||||||||||||||||||||||||-||||||||||||||||||||||||||||
Db     1 AGGGCAGAGGGGAGGAGGAAGAAG-AAAAGAGAGGGAGATAGTGTTATGACAAGAAAGACAGATTCAT 59

Qy  3491 tcaagggcagtgggaattgaccacaggggatatagtccacgtgatcctgggttctaggag 3550
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    60 TCAAGGGCAGTGGGAATTGACCACAGGGGATATAGTCCACGTGATCCTGGGTTCTAGGAG 119

Qy  3551 gcaaggctatattgtggggggaaaaaatcagttcaaggggagggagaacctgattct 3610
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 GCAGGGCTATATTGTGGGGGGAAAAAATCAGTTCAAGGGGAGGGAGAACCTGATTCT 179

Qy  3611 aatactatatattttcctttacaagctgagtcacaaggtagtaactg 3670
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   180 AATACTATATATTTTCCTTTACAAGCTGAGTAATTCACAAGGTAGTAACTG 239

Qy  3671 aggctgtaagattagttctctccttattggagactgtaaacaaaca 3730
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   240 AGGCTGTAAGATTAGTTCTCTCCTTATTAGGAGACTCTTTTTCTGTGTGGAGTTAGCA 299

Qy  3731 gcacaagggcaatcccgttctttttaacaggaagagaaaagcaaa 3790
         |||||||||||||||||||||||||||||||||||||||||||||
Db   300 GCACAAGGGCAATCCCGTTCTTTTTAACAGGAAGAGAAAAGCAAA 359

Qy  3791 cagattcaagcctaggtcttgctgctgtaatatgatctggttttcagcg 3850
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 CAGATTCAAGCCTAGGTCTTCGCTGCTATATATGATCTGGTTTTTGAAAAATCATTCAGCG 419

Qy  3851 atgtttactactctgattcagagaaaatgagactagtacccttt 3910
         ||||||||||||||||||||||||||||||||||||||||||||
Db   420 ATGTTTACTACTCTGATTCAGAGAAAATGAGACTAGTACCCTTTGGTCACCTGTAAACAAACA 479

Qy  3911 cccatttgtaaaatgtctcaacgttctttttcaggccttaac 3970
         ||||||||||||||||||||||||||||||||||||||||||
Db   480 CCCATTTGTAAAATGTCTCAAGTTCAGCGTTAACTGACAGAAACCAAATAAGAATAGAA 539

Qy  3971 tcttttagagcaaactgttgtttctctccactctgaggtgaa 4030
         ||||||||||||||||||||||||||||||||||||||||||
Db   540 TCTTTAGAGCAAACTGTTGTTTCTCCACTCTGAGGTGAGTCTCGAGGTGCAGTTTGGAAA 599

Qy  4031 tatttacttcacaagtattgacactgttgttggtattaac 4090
         ||||||||||||||||||||||||||||||||||||||||
Db   600 TATTTACTTCACAAGTATTGACACTGTTGTTGGTATTAACAACATAAAGTTGCTCAAAGG 659

Qy  4091 caatcattattccaagtgttaaagttactcctgcca 4150
         ||||||||||||||||||||||||||||||||||||
Db   660 CAATCATTATTTCAAGTGCCTTAAAGTTACTCGTGACAGTTTGGTTATATTTATTTGGCTA 719

Qy  4151 ttgccatttgcttttcgtttgggtttattaaatgg 4210
         |||||||||||||||||||||||||||||||||||
Db   720 TTGCCATTTGCTTTTTGTTTTTCTTTTGGGTTTATTAAATGTAAAGCAGGGATTATTAA 779

Qy  4211 cctacagtccagaaagcctgtgaatctgaggaaaaaatcg 4270
         ||||||||||||||||||||||||||||||||||||||||
Db   780 CCTACAGTCCAGAAAGCCTGTGAAATTGAATGAGGAAAAAATTACATTTTTTGTTTTACC 839

Qy  4271 accttctaactaaatttaacatttattcctttgcaaatagtgg 4330
         |||||||||||||||||||||||||||||||||||||||||||
Db   840 ACCTTCTAACTAAATTTAACATTTTATTCCATTGCGAATAGACCCATTAAACTCAAAAGTGG 899

Qy  4331 taataagtaactgtgattttgtcattaccaaatccccagacca 4390
         |||||||||||||||||||||||||||||||||||||||||||
Db   900 TAATAACAGTAACCTGTGTGATTTTGTCATTACCAAATCCCCAGACACATTTTATACTACTATA 959

Qy  4391 ttacagttgttgcagttacgttgtaggtgaaatattatacttccaaaaactacctttgaaatt 4450
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TTACAGTTGTTCCAGATACGTTGTAAGTGAAATATTTATACTCAAAAACTACTTTGAAATT 1019

Qy  4451 agacctcctgctggatcctgttcttaaacatattaataaacatgttcaaaatttgatat
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 AGACCTCCTGCTGGATCCTGTGTTCTTGTTTTTTAACTATAATAAAACATGTTTAAATTTGATAT 1079

Qy  4511 tttgataatcatattcattcattacatcgtttccttgttccttgtaactatactattatattga
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 TTTGATAATCATATTCATTCATTATCATCGTTTTCCTTTGTTTCCTTGTAACTATATTTATATTGA 1139

Qy  4571 aaacatcttctgagaagagttccccagaattcaccaatgagggttcttggcatgcacaca 4630
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 AAACATCTTTCTGAGAGAGTTCCCCAGATTTCGTTTCCTTTGTTTGTAATCTATATTTATATTTGA 1199

Qy  4631 cacagagtaagaacctgatttaggggctaaacaccttgacacattggtccctgaagactg 4690
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 CACAGAGTAAGAACTGATTTAGGAGGCTAAACATTGGTCCCTGAAGACTG 1259

Qy  4691 aaattaggaaagttctcccaaaataccacagttgtttcccaaaa 4750
         |||||||||||||||||||||||||||||||||||||||||
Db  1260 AAATTAGAAAGTTCTCCCAAAGATACACAGTTGTTTTTAAAGCTCAAGGTTCAGACGGGAA 1319

Qy  4751 tctgcccgcttctatagggaatgctctccctgggacctggtaggggtggtgtccttgtttct 4810
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1320 TCTGCCCGCTTCTATAGGAATGCTCTCCCCGGGACCTGGTAGGGTGCTGTCCTTGTTTCT 1379

Qy  4811 ggctggcttgtttatttctctgtcctagcgtctaaaggactgttttggatctccag 4870
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1380 GGCTGGCTTGTTTATTTTCTCTGTCCCTACGTGTTTAAAGGACTTGTTTTGGATCTCCAG 1439

Qy  4871 ttcctagcatagtgcccccgacagttctcaatgagttcgagagtgaatgaa 4930
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1440 TTCCTAGCATAGTGCCCTGCACAGTTGCAGGTGTTCTCAAGTGAGTGAATGGAA 1499

Qy  4931 tataaactagaaatatatcttcgttgttgaaatcagcacaccaccaccagtagtctcctggtt 4990
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500 TATAAAACTAGAAATATATCCCTGTTGAAAATCAGCACAGCCAGTTAGTCCCTGTCTCCTGGTG 1559

Qy  4991 tgtacgtcgtgtgtgtgtgtgtgtgtgtgtaaacaggtggaagtataggaa 5050
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1560 TGTAC----GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAGAGATATAGGAA 1615

Qy  5051 ctattatgggtatgggtgtcataaattgggatgttttttaaaagaaaaaactcccaaaca 5110
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1616 CTATTATTGGGATGGGTGCATAAATTGGGATGTTTTTTTAAAAGAAAAGTGAGAGACTGAAA 1675

Qy  5111 gacttctggaaggttgattctctaagaaatccagcacaccagtagtcctggttctaaggc 5170
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1676 GACTTCTGGAAGGTTGATTCTTTAAGAAATCTTGCTGGGCAGGTCGGAAGGCAACACCCCCTG 1735

Qy  5171 cacagcccccaccagcctcagtggccaacctctgtttctctcccccatgaagggctggctcc 5230
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1736 CACAGCCCCACCCAGCCTCACGTGTTCTCCCCCCATGAAGGGCTGGCTCC 1795

Qy  5231 cagtatatataaaacctctctgagctggggcatgaagcgaaggcaggggccaccatccaggca 5290
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1796 CAGTATATATAAAAACCTCTCTGGGACTCGGGCATGAGCCAGCAGCAGGCCACCATCCAGGCA 1855

Qy  5291 ccctctcagcacagc 5304
         |||||||||||||||
Db  1856 CCCTCTCAGCACAGC 1869
```

REFERENCE   1 (sites)
AUTHORS   Escribano,J., Ortego,J. and Coca-Prados,M.
TITLE   Isolation and characterization of cell-specific cDNA clones from a subtractive library of the ocular ciliary body of a single normal human donor: transcription and synthesis of plasma proteins
JOURNAL   J. Biochem. 118 (5), 921-931 (1995)
MEDLINE   96318503
REFERENCE   2 (sites)
AUTHORS   Stone,E.M., Fingert,J.H., Alward,W.L., Nguyen,T.D., Polansky,J.R., Sunden,S.L., Nishimura,D., Clark,A.F., Nystuen,A., Nichols,B.E., Ritch,R., Kalenak,J.W., Craven,E.R. and Sheffield,V.C.
TITLE   Identification of a gene that causes primary open angle glaucoma
JOURNAL   Science 275 (5300), 668-670 (1997)
MEDLINE   97158493
REFERENCE   3 (sites)
AUTHORS   Kubota,R., Noda,S., Wang,Y., Minoshima,S., Asakawa,S., Kudoh,J., Mashima,Y., Oguchi,Y. and Shimizu,N.
TITLE   A novel myosin-like protein (myocilin) expressed in the connecting cilium of the photoreceptor: molecular cloning, tissue expression, and chromosomal mapping
JOURNAL   Genomics 41 (3), 360-369 (1997)
MEDLINE   97312692
REFERENCE   4 (sites)
AUTHORS   Kubota,R., Kudoh,J., Mashima,Y., Asakawa,S., Minoshima,S., Hejtmancik,J.F., Oguchi,Y. and Shimizu,N.
TITLE   Genomic organization of the human myocilin gene (MYOC) responsible for primary open angle glaucoma (GLC1A)
JOURNAL   Biochem. Biophys. Res. Commun. 242 (2), 396-400 (1998)
MEDLINE   98113364
REFERENCE   5 (bases 1 to 1228)
AUTHORS   Shimizu,N. and Kudoh,J.
TITLE   Direct Submission
JOURNAL   Submitted (16-AUG-1997) to the DDBJ/EMBL/GenBank databases. Nobuyoshi Shimizu, Keio University School of Medicine, Department of Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan (E-mail:shimizu@dmb.med.keio.ac.jp, Tel:81-3-3351-2370(ex.2720), Fax:81-3-3351-2370)
FEATURES   Location/Qualifiers
     source   1..1228
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="FLEB14-14"
                     /cell_type="pre-pro-B cell"
                     /chromosome="1"
                     /clone="KB780C12"
                     /clone_lib="BAC"
                     /map="1q23-q24"
     misc_feature   216..221
                     /note="glucocorticoid responsive element consensus sequence"
     TATA_signal   371..375
     exon   431..1077
                     /number=1
BASE COUNT   312 a   325 c   350 g   241 t
ORIGIN

Query Match   7.9%; Score 417.8; DB 9; Length 1228;
Best Local Similarity 98.2%;
Matches 435; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

```
Qy  4862 gatcccaagtcctagcatagtgcctggacagtgcaggtgttccaatggagttgcagagt 4921
         ||||||||||||||||||||||||||||| ||||| ||||| |||| ||||| |||||||
Db     1 GATCTCCAGTTCCTAGCATAGTGCCTGGCACAGTGCCAGGTTCAATGAGTTTGCAGAGT   60

Qy  4922 gaatggaaatataaactaggaaatatatatccttgttgaaatgcaggcacacacagtagtcctggt 4981
         ||||||||||||||| |||||||||||||||||||||||||||||| |||| ||||| |||||||
Db    61 GAATGGAAATATAAACTAGGAAATATATATCCTTGTTGAAATGCAGCACACCACTAGTTCCTGGT  120

Qy  4982 gtaagtgtgtgtacgtgtgtgtgtgtgtgtgtgtgtaaaacccagtgtgag 5041
         |||||||||||||      |||||||||||||||||||||| |||||| |||
Db   121 GTAAGTGTGTGTAC------GTGTGTGTGTGTGTGTGTGTGTGTAAAACCAGTGTGGAG  174
```

RESULT 6
HSGLC1A1

LOCUS   HSGLC1A1   1086 bp   DNA   PRI   30-JAN-1998
DEFINITION Homo sapiens GLC1A (trabecular meshwork induced glucocorticoid response) gene, exon I, joined CDS.
ACCESSION   Z97171
VERSION   Z97171.1 GI:2425156
KEYWORDS   GLC1A.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1086)
AUTHORS   Stone,E.M., Fingert,J.H., Alward,W.L., Nguyen,T.D., Polansky,J.R., Sunden,S.L., Nishimura,D., Clark,A.F., Nystuen,A., Nichols,B.E., Ritch,R., Kalenak,J.W., Craven,E.R. and Sheffield,V.C.
TITLE   Identification of a gene that causes primary open angle glaucoma
JOURNAL   Science 275 (5300), 668-670 (1997)
MEDLINE   97158493
REFERENCE   2 (bases 1 to 1086)
AUTHORS   Adam,M.F., Belmouden,A., Binisti,P., Brezin,A.P., Valtot,F., Bechetoille,A., Dascotte,J.C., Copin,B., Gomez,L., Chaventre,A., Bach,J.F. and Garchon,H.J.
TITLE   Recurrent mutations in a single exon encoding the evolutionarily conserved olfactomedin-homology domain of TIGR in familial open-angle glaucoma
JOURNAL   Hum. Mol. Genet. 6 (12), 2091-2097 (1997)
MEDLINE   97424261
REFERENCE   3 (bases 1 to 1086)
AUTHORS   Garchon,H.J.
TITLE   Direct Submission
JOURNAL   Submitted (20-JUN-1997) Garchon H.-J., Hopital Necker, INSERM U25, 161 rue de Sevres 75743 Paris cedex 15 FRANCE
FEATURES   Location/Qualifiers
     source   1..1086
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="q23-q25"
                     /tissue_type="white blood cells"
     TATA_signal   355..363
     exon   355..1065
                     /gene="GLC1A"
                     /label=ex1
                     /number=1
     mRNA   join(385..1065,297177.1:21..146,297174.1:21..1475)
                     /gene="GLC1A"
                     /label=RNAGLA1A
     gene   join(385..1065,297177.1:21..146,297174.1:21..1475)
                     /gene="GLC1A"
     misc_feature   462..1065

```
Qy  5042 atataggaactatattatggggtgtatggggttgcataaattggatgttcatttttaaaaagaaa 5101
         ||||||||||||| ||||||||| ||||||| |||||||| ||||||||||| ||||||||||||
Db   175 ATATAGGAACTATTATTATTGGGGTATGGGGTATGGGTGCATAAAATTGGGATGTTCTTTTTAAAAGAAA 234

Qy  5102 ctccaaacagacttctggaagttcatttctaagaaatccttctggaagccagcgtgaaggcaac 5161
         ||||||||||||||| |||||| ||||||||||||| |||| |||| ||| || |||||||||||
Db   235 CTCCAAACAGACTTCCGGAAGGTTATTTCTAAGAATCTTCTGGCAGCCGTGAAGGCAAC       294

Qy  5162 ccccctgtgcacaggccacccacccaccagcctcacgtggccacctctggtcctcccccatgaaggg 5221
         |||||||||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db   295 CCCCCTGTGCACAGGCCACCCCCACCCACCCAGCCTCACGTGGCCACCTCTGGTCCTCCCCCATGAAGGG 354

Qy  5222 ctgggctccccagtatatataaactctctggaactctggggccatgtgataagggcacc 5281
         |||||||||||||||||||||| ||||||||||  |||| |||||| |||| |||||||
Db   355 CTGGGCTCCCCAGTATATATAAAACTCTCTGGAGGCTCGGGCCATGAGCCAGCAGCAAGGGCACCC 414

Qy  5282 atccaggcacctctcagcacagc 5304
         |||||||||||||||||||||||
Db   415 ATCCAGGCACCTCTCAGCACAGC  437
```

```
RESULT   7
LOCUS       AC007688/c
LOCUS       AC007688      161577 bp      DNA      PRI      01-SEP-1999
DEFINITION  Homo sapiens 12p12-27.2-31.7 BAC RPCI11-392P7 (Roswell Park Cancer
            Institute Human BAC Library) complete sequence.
ACCESSION   AC007688
VERSION     AC007688.15  GI:5815499
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161577)
```

```
AUTHORS   Muzny,D., Arenson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,
          Dugan,S., Durbin,J., Forcum,J., Garcia,C., Gorrell,J.H.,
          Gorrell,L.L., Hernandez,J., Hernandez,J., Issar,A., Jackson,L., Kneitz,S.,
          Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichtarge,O., Liu,W.,
          Logan,O., Lu,J., Marondel,I., Martinez,C., Merscher,S., Miller,A.,
          Montgomery,K., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,
          Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,
          Vo,Q., Williamson,A., Worley,K.C., Xhang,A.M., Yang,R., Yu,W.,
          Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.
```

```
TITLE     Direct Submission
JOURNAL   Unpublished
REFERENCE 2  (bases 1 to 161577)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3  (bases 1 to 161577)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Sep 1, 1999 this sequence version replaced gi:5757565.
  INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu
COMMENT     CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
            STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
            Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
            Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25.3389:3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.
            html.
```

```
----------- QUALSTAT-REPORT -----------

----------- Summary Statistics -----------
Contig length:                                                    161577
Phrap values in estimate:                                         160751
Average error rate (BCM-Phrap estimate):                        0.000163681
Fraction of Phrap values less than 40:                           0.0376047
Number of consensus changing edits:                                30
Number of N's in consensus:                                         0

----------- Consensus changing edits -----------
```

```
BASE COUNT     277 a      282 c      316 g      211 t
ORIGIN
```

CDS

```
/gene="GLC1A"
/note="first coding exon"
/label=start
join(462..1065,297177.1.21..146,297174.1.21..805)
/gene="GLC1A"
/codon_start=1
/label=CDSGLA1A
/protein_id="CAB09899.1"
/label=CDSGLA1A
/db_xref="GI:2425157"
/db_xref="SWISS-PROT:Q99972"
/translation="MRFFCARCCSFGPEMPAVQLLLACLVWDVGARTAQLRKANDOS
GRCQTTFSVASPNESSCPEQSQAMSVIHNLQRDSSTORDLEATKARLSSLESLLHQL
TLDDAARPQETQELQRELGTLRRERDQLETOTRELETAYSNLRDKSVLEEEKKRLR
OENENLARRLEESSQEVARLRRGOCPQTRDTARAVPPGSREVSTWNLDTLAFQELKSE
LIEVPASRILKESPSGYLRSGEGDTCGCELWVGEPLTLRTAETTIGKYGVWMRDKP
TYPYTQETWRIDTVGTDVRQVFEYDLISQFMQYPSKVHILPRPLESTGAVVSGSL
YFQGAESRVIVRYELNTETVKAEKEIPGACYHGQFPYSWGGYTDIDLAVDEAGLWVIY
STDEAKGAIVLSKLNPENLELEOTWETNIRKOSVANAFIICGTLYTVSSYSADATVN
FAYDTGTGISKTLIIPFKNRYKYSSMIDYNPLEKKLFAWDNLNMVTYDIRLSKM"
```

```
Qy  4876  agcatagtagttgcctggcacagcgcaggttctcaatgagtttgcagagtgaatggaatata  4935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AGCATAGTGCCTGGCACAGTGCAGGTTCTCAATGAGTTTGCAGAGTGAAATATAA  60

Qy  4936  actagaaatatatcttttgttgaaatccagcacgcaccccagtagtcctgtggttgtgttgtac  4995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  ACTACAGAAATATATCCTTTGTTCGAAAATCAGCACACCAGTAGTCCTCTGGGTGTAAGTGTGTAC  120

Qy  4996  gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgagatataggaactatt  5055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  -----GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAAAACCAGGTGGGAGATATAGGAACTATT  176

Qy  5056  attgggtatgggtgcataaattggatgtctcttttaaaaaagaaaactccaaacagactt  5115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   177  ATTGGGTATGGGTGCATAAATTGGGATGTCTCTTTTAAAAAGAAAACTCCAAACAGACTT  236

Qy  5116  ctggaaggttattctcaagaatcttgctgctggcagcgtgaaggcaacccccctgtgcacag  5175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237  CTGGAAGGTTATTTCTAAGAAATCTTGCTGCCTGAAGGCAACCCCCTGTGCACAG  296

Qy  5176  cccccaccaggccaccctctgtcctcccccatgaaggggctgtgctcccccagta  5235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297  CCCACCACCAGCTCACGTGGCCAACCCTCTGTCTCCCCCATGAAGGGCGTGGCTCCCCCAGTA  356

Qy  5236  tatataaaactcctctgcgggagccggaggcagggccaccccaatccagggaacctct  5295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357  TATATAAACCTCTCTGGAGCTCGGAGCCGGGCATGAGCCACCAAGGCCACCCATCCAGGCAACCTCT  416

Qy  5296  cagcacagc  5304
          |||||||||
Db   417  CAGCACAGC  425
```

```
Query Match         7.7%;      Score 407.8;  DB 66;  Length 1086;
Best Local Similarity  98.6%;   Pred. No. 1.7e-79;
Matches 423; Conservative    0; Mismatches   2;  Indels   4;  Gaps   1;
```

Position	Original+Context	Edited+Context
7033	acctcgccgt(n)ccgccccctt	acctcgccgt(c)ccgccccctt
47567	aaaaaaaaaa(n)ggaaaataat	aaaaaaaaaa(a)ggaaaataat
51135	aaagaaagaa(n)aaagaaggaa	aaagaaagaa(a)aaagaaggaa
75582	aaaaaaaaaa(n)aaangcatgtt	aaaaaaaaaa(a)aaangcatgtt
75585	aaaaaaaaaa(n)ggaggttcgc	aaaaaaaaaa(a)ggaggttcgc
75667	ctaaagcaga(n)taagatctta	ctaaagcaga(a)taagatctta
75752	ttttaaatag(n)gcttgctt	ttttaaatag(t)gcttgctt
84017	gggaggag(n)aaggaaggag	gggaggag(a)aaggaaggag
85227	tttgtttt(n)tgtttttttt	tttgtttt(t)tgtttttttt
99681	gggttcaag(n)atgcagtgag	gggttcaag(a)atgcagtgag
111307	aatctcttat(n)atcntttttt	aatctcttat(g)atcntttttt
111374	gtatacnatc(n)tctttttttt	gtatacnatc(t)tctttttttt
111378	gtatacatc(n)tctttttttt	gtatacatc(t)tctttttttt
112621	aaaaaaaaaa(n)ccatcctaga	aaaaaaaaaa(a)ccatcctaga
135812	attcacctc(c)tttttttttt	attcacctc(c)tttttttttt
137207	ttgcagcac(n)cgccaccacg	ttgcagcac(a)cgccaccacg
137218	cgccaccacg(n)ctggctaagt	cgccaccacg(a)ctggctaagt
145113	gcaagtgaaa(n)natgtagaat	gcaagtgaaa(a)natgtagaat
145114	caagtgaaan(n)atgtagaat	caagtgaaan(g)atgtagaat
145232	ggcaccggtg(n)ctcacccagt	ggcaccggtg(c)ctcacccagt
145233	gcaccggtgn(n)tcacccagt	gcaccggtgn(t)tcacccagt
145238	ggtgnntcac(n)ccagtaatcc	ggtgnntcac(c)ccagtaatcc
145321	gcaacatggt(t)nnncccatc	gcaacatggt(t)nnncccatc
145322	gcaacatggt(n)nnccccatc	gcaacatggt(n)nnccccatc
145323	caacatggtt(n)nccccatctc	caacatggtt(n)nccccatctc
145324	aacatggttn(n)cccatctc	aacatggttn(c)cccatctc
145363	gttcaccagg(n)gtgttggcgt	gttcaccagg(c)gtgttggcgt
145377	gtggcgtgta(t)ttgtagtcct	gtggcgtgta(t)ttgtagtcct
145397	tagctactg(n)gaggaggat	tagctactg(g)gaggaggat
145505	aaaaaaaaaa(n)gaaaaaaa	aaaaaaaaaa(a)gaaaaaaa

------- Distribution of Quality < 40 Bases -------

```
       1000|
        900|
        800|
        700|
   #    600|
 bases  500|                    *   *   *
        400|                *   *   *   *   *
        300|            *   *   *   *   *   *   *
        200|        *   *   *   *   *   *   *   *   *
        100|    *   *   *   *   *   *   *   *   *   *
          0|*   *   *   *   *   *   *   *   *   *   *
            5   10  15  20  25  30  35  40
                  Phrap Value Range
```

FEATURES

Version: 1.01 gxfo.
Location/Qualifiers

FEATURES		
source	1..161577	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="12p13.3"	
	/clone="RPCI11-392P7"	
repeat_region	complement(3..65)	
	/rpt_family="MER4B"	
repeat_region	161..471	
	/rpt_family="AluSx"	
repeat_region	complement(607..702)	
	/rpt_family="MIR"	
repeat_region	703..838	
	/rpt_family="FLAM_C"	
repeat_region	complement(839..916)	
	/rpt_family="MIR"	
repeat_region	1205..1377	
	/rpt_family="MER104"	
STS	1388..1525	
	/standard_name="WIAF-759-STS"	
repeat_region	1683..1971	
	/rpt_family="MIR"	
repeat_region	1974..2261	
	/rpt_family="AluJo"	
repeat_region	2486..2593	
	/rpt_family="AluJb"	
repeat_region	2610..3214	
	/rpt_family="MIR"	
misc_feature	/note="Region: Unigene cluster similar to AA056332 and	
	AI074576."	
	/db_xref="dbSTS:65358"	
STS	2712..2823	
	/standard_name="SHGC-44583"	
	/db_xref="dbSTS:48451"	
repeat_region	3185..3491	
	/rpt_family="AluSq"	
gene	U68487. 4509..5919	
	/gene="Human 5-hydroxytryptamine7 receptor isoform b mRNA,	
repeat_region	6967..7158	
	/rpt_family="(CCCCG)n"	
repeat_region	7094..7240	
	/rpt_family="AluSq"	
repeat_region	complement(6397..6537)	
	/rpt_family="L2"	
gene	AI380369. join(7241..7357,17979..18117,20062..20242,23559..23933)	
	/gene="Unigene cluster containing N42239, R09325, and	
repeat_region	complement(7770..8250)	
	/rpt_family="(CCG)n"	
repeat_region	8831..9073	
	/rpt_family="MER44B"	
repeat_region	complement(9075..9303)	
	/rpt_family="MER102"	
repeat_region	9671..10135	
	/rpt_family="L2"	
repeat_region	10426..10592	
	/rpt_family="LTR33A"	
repeat_region	10625..10689	
	/rpt_family="MIR"	
repeat_region	11094..11322	
	/rpt_family="MIR"	
repeat_region	complement(11985..12403)	
	/rpt_family="L2"	

Query Match

Best Local Similarity 78.5%;
Matches 227; Conservative 0; Score 189.8; DB 10; Length 161577;
Pred. No. 1.3e-31; Mismatches 62; Indels 0; Gaps 0;

```
Qy 1280 agggtgagggtctggtgtcttacacctacctgtatgctctcagcctccccgcgtagctgggactacag 1399
        |||||||  ||| |||||||||||||||| ||||||||||||||||  |||||||||||||||||||||
Db 29160 AGAGTCTCACTCTATTACCCAGGGCTGAAGTGCAGTGGCGGGCGCCAAATCTCGGCTCACTGCAACC 29101

Qy 1340 tctggcctccaggtcaaggcaattctcctgtctccagcctacacctgtatgctctcagcctccccgcgtagctgggactacag 1399
Db 29100 TCTGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAG 29041

Qy 1400 gcgcaacgcccggctaaatttgtattgttagtagagatggggtttcaccatattagcccg 1459
        || ||| |||||| ||||   |||| |||||| |||||||||||||| |||||||||||||
Db 29040 GCACACGCCCAGCTAATTTTTGTAATTTTAGTACAGATGGGGTTTTCACCATATTGGCCAG 28981

Qy 1460 gctggtcttgaactcctgacctcaggtgatccacccaccccagcctcctaaagtgctggg 1519
        ||||||||||||||||||||||| || ||| ||||||||||||||||||||||||||||||
Db 28980 GCTGGTCTCAAAACTCCTGACCTCAGGTGATCCACCCGCCTCGCCTCCCAAAGTGCTGGG 28921

Qy 1520 attacaggcatgagtcaccgcgcccggccaaggtcagtgtttaataag 1568
        ||||||||| |||||||||||||||||| ||||  |||||||| || ||
Db 28920 ATTACAGGTGTGAGCCACCACGCCCGGCCTGGAAAGGGTTTTAAAGTAG 28872
```

RESULT 8
AC023790/c

LOCUS   AC023790   193123 bp   DNA   HTG   25-JUL-2000
DEFINITION   Homo sapiens chromosome 12 clone RP11-377D9, WORKING DRAFT

```
SEQUENCE, 35 unordered pieces.
ACCESSION   AC023790
VERSION     AC023790.16  GI:9438256
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193123)
AUTHORS     Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 193123)
AUTHORS     Worley,K.C.
TITLE       Direct Submission
JOURNAL     Submitted (18-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 25, 2000 this sequence version replaced gi:9255941.
            --------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information
            Center project name: HABM
            Center clone name: RP11-377D9
            --------------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 167304 bases at least Q40
            Consensus quality: 176244 bases at least Q30
            Consensus quality: 181341 bases at least Q20
            Consensus quality: 181341 bases at least Q20
            Estimated insert size: 180060; sum-of-contigs estimation
            Estimated insert size: 180060; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-fp estimation
            Quality coverage: 3.1x in Q20 bases; sum-of-contigs estimation
            --------------------
            * NOTE: Estimated insert size may differ from sequence length
            * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 35 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *      1: contig of 21094 bp in length
            *  21095: gap of unknown length
            *  21194: contig of 21094 bp in length
            *  33195: gap of 12736 bp in length
            *  33931: contig of unknown length
            *  34030: gap of unknown length
            *  34031: contig of 14288 bp in length
            *  48318: gap of unknown length
            *  48418: contig of unknown length
            *  48819: gap of 10381 bp in length
            *  58799: contig of unknown length
            *  58899: gap of unknown length
            *  58900: contig of 9824 bp in length
```

```
            *  68724: gap of unknown length
            *  68824: contig of 7250 bp in length
            *  76074: gap of unknown length
            *  76174: contig of 6695 bp in length
            *  82868: gap of unknown length
            *  82869: contig of 7078 bp in length
            *  82968: gap of unknown length
            *  82969: contig of 7078 bp in length
            *  90046: gap of unknown length
            *  90047: contig of 5995 bp in length
            *  90146: gap of unknown length
            *  90147: contig of 5995 bp in length
            *  96141: gap of unknown length
            *  96142: contig of 5632 bp in length
            *  96241: gap of unknown length
            *  96242: contig of 5632 bp in length
            * 101873: gap of unknown length
            * 101874: contig of 6029 bp in length
            * 101973: gap of unknown length
            * 101974: contig of 6029 bp in length
            * 108002: gap of unknown length
            * 108003: contig of 4801 bp in length
            * 108102: gap of unknown length
            * 108103: contig of 4801 bp in length
            * 112903: gap of unknown length
            * 112904: contig of 7206 bp in length
            * 113003: gap of unknown length
            * 113004: contig of 7206 bp in length
            * 120209: gap of unknown length
            * 120210: contig of 6044 bp in length
            * 120309: gap of unknown length
            * 120310: contig of 6044 bp in length
            * 126353: gap of unknown length
            * 126354: contig of 6435 bp in length
            * 126453: gap of unknown length
            * 126454: contig of 6435 bp in length
            * 132888: gap of unknown length
            * 132889: contig of 7149 bp in length
            * 132988: gap of unknown length
            * 132989: contig of 7149 bp in length
            * 140137: gap of unknown length
            * 140138: contig of 4394 bp in length
            * 140237: gap of unknown length
            * 140238: contig of 4394 bp in length
            * 144631: gap of unknown length
            * 144632: contig of 5581 bp in length
            * 144731: gap of unknown length
            * 144732: contig of 5581 bp in length
            * 150312: gap of unknown length
            * 150313: contig of 3702 bp in length
            * 150412: gap of unknown length
            * 150413: contig of 3702 bp in length
            * 154114: gap of unknown length
            * 154115: contig of 4244 bp in length
            * 154214: gap of unknown length
            * 154215: contig of 4244 bp in length
            * 158458: gap of unknown length
            * 158459: contig of 2827 bp in length
            * 158558: gap of unknown length
            * 158559: contig of 2827 bp in length
            * 161385: gap of unknown length
            * 161386: contig of 2835 bp in length
            * 161485: gap of unknown length
            * 161486: contig of 2835 bp in length
            * 164320: gap of unknown length
            * 164321: contig of 3294 bp in length
            * 164420: gap of unknown length
            * 164421: contig of 3294 bp in length
            * 167714: gap of unknown length
            * 167715: contig of 3200 bp in length
            * 167814: gap of unknown length
            * 167815: contig of 3200 bp in length
            * 171014: gap of unknown length
            * 171015: contig of 2729 bp in length
            * 171114: gap of unknown length
            * 171115: contig of 2729 bp in length
            * 173843: gap of unknown length
            * 173844: contig of 2688 bp in length
            * 173943: gap of unknown length
            * 173944: contig of 2688 bp in length
            * 176631: gap of unknown length
            * 176632: contig of 1917 bp in length
            * 176731: gap of unknown length
            * 176732: contig of 1917 bp in length
            * 178648: gap of unknown length
            * 178649: contig of 1484 bp in length
            * 178748: gap of unknown length
            * 178749: contig of 1484 bp in length
            * 180232: gap of unknown length
            * 180233: contig of 2772 bp in length
            * 180332: gap of unknown length
            * 180333: contig of 2772 bp in length
            * 183104: gap of unknown length
            * 183105: contig of 1516 bp in length
            * 183204: gap of unknown length
            * 183205: contig of 1516 bp in length
            * 184720: gap of unknown length
            * 184721: contig of 1317 bp in length
            * 184820: gap of unknown length
            * 184821: contig of 1317 bp in length
            * 186137: gap of unknown length
            * 186138: contig of 1976 bp in length
            * 186237: gap of unknown length
            * 186238: contig of 1976 bp in length
            * 188213: gap of unknown length
            * 188214: contig of 1760 bp in length
            * 188313: gap of unknown length
            * 188314: contig of 1760 bp in length
            * 190073: gap of unknown length
            * 190074: contig of 1682 bp in length
            * 190173: gap of unknown length
            * 190174: contig of 1682 bp in length
            * 191855: gap of unknown length
            * 191856: contig of 1168 bp in length
            * 191955: gap of unknown length
            * 191956: contig of 1168 bp in length
FEATURES             Location/Qualifiers
     source          1..193123
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-377D9"
                     /chromosome="12"
BASE COUNT    54294 a   41007 c   41291 g   53110 t    3421 others
ORIGIN
```

Query Match   3.6%;   Score 189.8;   DB 77;   Length 193123;
Best Local Similarity 78.5%;
Matches 227;   Conservative 0;   Mismatches 62;   Indels 0;   Gaps 0;
Pred. No. 1.3e-31;

```
Qy   1280 agggtgaggtctgtgtcttacacctgtatgctcctacacctgagctcactgcaacc 1339
          |||||  ||| || |  || ||| ||  |||    || |  ||| ||| |||||||
Db 134698 AGAGTCTCACTCTATTACCCAGGCTGAAGTGCGGCGAACTCGGCTCATTGCAACC 134639

Qy   1340 tctgcctcccaggttcaagcaattctcctgtcccgcgtagctgggactacag 1399
          |||||||||||||||||||  |||||||||| || |  ||||||||||||||
Db 134638 TCTGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAG 134579

Qy   1400 gcgcagcccggctaatttttt gtatt... 1459
          ||| || || |||||||| || |||||
Db 134578 GCACACGCCAGGCTAATTTTGTATTTTTAGTAGAGATAGGGGTTTCACCATATTAGCCAG 134519

Qy   1460 gctgggtcttgaacctcctgacctcaggtgat... 1519
          ||||||||| ||| |||||||||||||||||
Db 134518 GCTGGGTCTTCAAACTCCTGACCTCAGGTGATCAACCCTGCCTCAGCCTCCCTAAAGTGCTGGG 134459

Qy   1520 attacaggcatgagtcaccgcgcccggccaagggtcagtgttaataag 1568
          |||||| | |||| || |||| ||| ||||     ||||| |||||
Db 134458 ATTACAGGTGTGAGCCACCACGCCCGGCCTGAAAGGGTTTTAAAGTAG 134410
```

RESULT 9

```
LOCUS       AC012404   199722 bp   DNA      HTG   26-MAY-2000  WORKING DRAFT
DEFINITION  Homo sapiens chromosome 15 clone RP11-39M21 map 15, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC012404
VERSION     AC012404.4  GI:7107808
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 199722)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-39M21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199722)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkly,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,J., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,S., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 28, 2000 this sequence version replaced gi:6514012.
            All repeats were identified using RepeatMasker:
            Smit,A.F.A. & Green,P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------------  Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----------------  Project Information
            Center project name: L2261
            Center clone name: 39_M_21
            ----------------  Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 178894 bases at least Q40
```

```
            Consensus quality: 191626 bases at least Q30
            Consensus quality: 196426 bases at least Q20
            Insert size: 199722; agarose-fp
            Insert size: 199222; sum-of-contigs
            Quality coverage: 6.2 in Q20 bases; agarose-fp
            Quality coverage: 6.1 in Q20 bases; sum-of-contigs
            ---------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            ---------------.
            *            1           5729: contig of 5729 bp in length
            *         5730           5829: gap of 100 bp
            *         5830          20896: contig of 15067 bp in length
            *        20897          20996: gap of 100 bp
            *        20997          52603: contig of 31607 bp in length
            *        52604          52703: gap of 100 bp
            *        52704          89348: contig of 36645 bp in length
            *        89349          89448: gap of 100 bp
            *        89449         138760: contig of 49312 bp in length
            *       138761         138860: gap of 100 bp
            *       138861         199722: contig of 60862 bp in length.
FEATURES             Location/Qualifiers
     source          1..199722
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /clone="RP11-39M21"
                     /clone_lib="RPCI-11 Human Male BAC"
                     /map="15"
     misc_feature    1..5729
                     /note="assembly_fragment"
     misc_feature    5830..20896
                     /note="assembly_fragment"
     misc_feature    20997..52603
                     /note="assembly_fragment"
     misc_feature    52704..89348
                     /note="assembly_fragment"
                     vector_side:left
     misc_feature    89449..138760
                     /note="assembly_fragment"
                     vector_side:left
     misc_feature    138861..199722
                     /note="assembly_fragment"
                     clone_end:SP6
                     vector_side:right
BASE COUNT    60442 a  42775 c  40458 g  55547 t  500 others
ORIGIN
```

```
Query Match          3.6%;  Score 189;  DB 71;  Length 199722;
Best Local Similarity  80.1%;  Pred. No. 1.9e-31;
  Matches 222;  Conservative  0;  Mismatches 55;  Indels  0;  Gaps  0;

Qy   1290 tctgtgtcttacacctacctgagctcaacctgcaacctgcctccc 1349
          |||||  ||  || || || |||||| ||||| |  |  |||||
Db 184619 TCTGTCGCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCCCTGGCAACCTCTGCCTCCC 184678

Qy   1350 aggttcaagcaattctcctgtctcagcctcccgtagctgggactacaggcgcc 1409
          |||||| || | ||||||| ||||||||||| |||||||||| ||| | |
Db 184679 AGGTTCGGGTGGATTCTCCTCCCTCAGCCTCCCAAGTAGCTGGGATTACAAGTGCACACT 184738

Qy   1410 ggctaattttttgtattgttagtagagatggggtttcaccatattagcccggctggtcttg 1469
          ||||||||||||||||| |||||||||| |||||||||||||||| |||| |||||||||
Db 184739 GGCTAATTTTTTGTATTTTTAGTAGAGAAAGGGTTTCACCATATTGGCCAGGCTGGTCTTG 184798

Qy   1470 aacttcctgacctcaggtgatccacccacctcagcctccc...ggca 1529
          ||| ||||||||||||||||||||||||||||||||||||   |||
Db 184799 AACTCCTGACCTCAGGTGATCCACCCACCTCAGCCTCCCAAAAGTGCTGGGATTACAGGCG 184858
```

```
Qy   1530 tgagtcaccgcccgcccggccaaggggtcagtgttaata 1566
          ||||·|||| ||| |||·||| ||| ||||·|| |·||||·|
Db 184859 TGAGCCACCATGCCAGGCCATTCTTACTTTTTACA 184895
```

```
RESULT 10
LOCUS       AC004973/c
DEFINITION  AC004973          97037 bp      DNA            PRI      21-DEC-1999
ACCESSION   Homo sapiens PAC clone RP5-1139I1 from Xq23, complete sequence.
VERSION     AC004973
KEYWORDS    AC004973.1  GI:3694660
SOURCE      HTG.
            human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 97037)
AUTHORS     Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
REFERENCE   2  (bases 1 to 97037)
AUTHORS     Courtney,L., Langston,Y. and Drone,K.
TITLE       The sequence of Homo sapiens PAC clone RP5-1139I1
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 97037)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (03-OCT-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAR-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 3, 1998 this sequence version replaced gi:3213024.
            -------------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
            Center project name: H_DJ1139I01
            --------------
            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.
            NOTICE: This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.
            MAPPING INFORMATION:
            This sequence was generated from part of bacterial clone contigs of
            human chromosome X, constructed by the chromosome X mapping group
```

at the Sanger Centre, Wellcome Trust Genome Campus, Hinxton, UK.
Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX/

SOURCE INFORMATION:
This clone was derived from human PAC library RPCI-5, prepared by
Pieter de Jong and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu) using the method described by
Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from
one male donor.
The clone may be obtained either from Genome Systems, Inc.
(http://www.genomesystems.com) or Research Genetics, Inc.
(http://www.resgen.com); or from Pieter de Jong.
VECTOR: pCYPAC2
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP4-555N2; the clone sequenced
to the right is RP3-404F18, 200 bp overlap. Actual start of this
clone is at base position 1 of RP5-1139I1; actual end is at 13269
of RP3-404F18.

```
FEATURES             Location/Qualifiers
     source          1..97037
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq23"
                     /clone="RP5-1139I1"
                     /clone_lib="RPCI-5"
     repeat_region   530..801
                     /rpt_family="L2"
     repeat_region   830..1035
                     /rpt_family="Alu"
     repeat_region   2179..2349
                     /rpt_family="MER1_type?"
     repeat_region   3434..3489
                     /rpt_family="MIR"
     repeat_region   3526..3739
                     /rpt_family="MIR"
     repeat_region   4681..4852
                     /rpt_family="MIR"
     repeat_region   5395..5961
                     /rpt_family="MER4-group"
     repeat_region   6203..6393
                     /rpt_family="L1"
     repeat_region   6394..7105
                     /rpt_family="L1"
     repeat_region   7104..7211
                     /rpt_family="L1"
     repeat_region   7342..7520
                     /rpt_family="L1"
     repeat_region   7542..7644
                     /rpt_family="Alu"
     repeat_region   7645..8674
                     /rpt_family="Retroviral"
     repeat_region   8682..9314
                     /rpt_family="Retroviral"
     repeat_region   9315..9368
                     /rpt_family="Retroviral"
     repeat_region   9371..10157
                     /rpt_family="7SLRNA"
     repeat_region   10165..10817
                     /rpt_family="Retroviral"
     repeat_region   10818..10973
                     /rpt_family="Retroviral"
     repeat_region   10974..11111
                     /rpt_family="Alu"
     repeat_region   11238..11612
                     /rpt_family="L1"
     repeat_region   11972..12283
                     /rpt_family="MaLR"
     repeat_region   12570..12854
                     /rpt_family="Alu"
     repeat_region   12859..13433
                     /rpt_family="L1"
     repeat_region             ..
                     /rpt_family="MER1_type"
```

```
repeat_region    13435..13564
                 /rpt_family="Alu"
repeat_region    13580..13845
                 /rpt_family="MER1_type"
repeat_region    13850..14118
                 /rpt_family="L1"
repeat_region    14119..14411
                 /rpt_family="Alu"
repeat_region    14412..14813
                 /rpt_family="L1"
repeat_region    14801..15447
                 /rpt_family="L1"
repeat_region    15501..16654
                 /rpt_family="L2"
repeat_region    17412..17533
                 /rpt_family="MER1_type"
repeat_region    17534..17793
                 /rpt_family="MER1_type"
repeat_region    17794..17912
                 /rpt_family="Alu"
repeat_region    17793..18249
                 /rpt_family="MER1_type"
repeat_region    18293..18358
                 /rpt_family="L1"
repeat_region    18348..18476
                 /rpt_family="MIR"
repeat_region    18483..18908
                 /rpt_family="L2"
repeat_region    18932..19227
                 /rpt_family="MaLR"
repeat_region    19390..19604
                 /rpt_family="Alu"
repeat_region    19606..20227
                 /rpt_family="L2"
repeat_region    20228..20477
                 /rpt_family="AcHobo"
repeat_region    20478..20721
                 /rpt_family="Alu"
repeat_region    20738..20828
                 /rpt_family="AcHobo"
repeat_region    20848..21707
                 /rpt_family="AcHobo"
repeat_region    21717..22111
                 /rpt_family="L2"
repeat_region    22163..22378
                 /rpt_family="L2"
repeat_region    22866..23031
                 /rpt_family="L1"
repeat_region    23060..23339
                 /rpt_family="L1"
repeat_region    23409..24201
                 /rpt_family="Alu"
repeat_region    24403..24554
                 /rpt_family="L2"
repeat_region    24420..24554
                 /rpt_family="AcHobo"
repeat_region    24572..25105
                 /rpt_family="AcHobo"
repeat_region    25106..25479
                 /rpt_family="L1"
repeat_region    25480..25768
                 /rpt_family="Alu"
repeat_region    25769..25985
                 /rpt_family="L1"
```

```
RESULT 11
LOCUS       HUMYWXD703/c
DEFINITION  HUMYWXD703 135038 bp      DNA          PRI      24-DEC-1996
ACCESSION   Homo sapiens ADP/ATP carrier protein (ANT-2) gene, complete cds.
VERSION     L78810
KEYWORDS    L78810.1  GI:1381111
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135038)
  AUTHORS   Chen,C.N., Su,Y., Baybayan,P., Siruno,A., Nagaraja,R.,
            Mazzarella,R., Schlessinger,D. and Chen,E.
  TITLE     Ordered shotgun sequencing of a 135 kb Xq25 YAC containing ANT2 and
            four possible genes, including three confirmed by EST matches
  JOURNAL   Nucleic Acids Res. 24 (20), 4034-4041 (1996)
  MEDLINE   97078684
COMMENT     Submitted by:
            Elson Chen,
            Advanced Center for Genetic Technology,
            Applied Biosystems Division of Perkin Elmer Corp., 850 Lincoln
            Center Drive,
            Foster City, CA 94404 USA
            and
            David Schlessinger,
            Department of Molecular Microbiology and Center for Genetics in
            Medicine,
            Washington University School of Medicine,
            St. Louis MO 63110 USA
            e-mail: ellson@genseq.apldbio.com and davids@genetics.wustl.edu
            Note: Gene predictions were accomplished with runs of Grail
            versions 1.1 and 1.2, coupled with fasta and blastx comparisons to
            genbank & non-redundant peptide libraries. Repeat analysis was
            accomplished via censor.
            The Rat EST105369 shows significant homology via blastx to this
            sequence
            Strand Start     End
            top    9072      9326
            The Graves Disease carrier protein (X60035) shows significant
            homology via blastx to this sequence
            Strand Start End
            top    28802     29076
            top    35858     36081
            Comments for gene ANT-2 :
            This gene shows homology via blastx to the EST clone yl15a12 Strand
            Start End
            bottom 125256 125915.
FEATURES             Location/Qualifiers
     source          1..135038
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="X"
     repeat_unit     complement(271..560)
```

```
Db  54862 TCTGCCTCCTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGTGGACTACAG 54803

Qy  1400  gcg-------cacgcccggctaattttgtattgttagtgtagagatgggtttcaccata 1451
          |||       ||||||  | |||||| ||  ||||||||||||||||||| |||||||||
Db  54802 GCGCAATGCCACCACGCCCACGTAATTTTGTATTTTTAGTAGAGGATTAGAATGGGGTTTCACCATA 54743

Qy  1452  ttagcccggctggtcttgaactcctgacctcaggtgatccaccaacctcctaa 1511
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  54742 TTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCACCAACCTCCTAA 54683

Qy  1512  gtgtctgggattacaggcgtgagtcaccgcgcccggccaaggtcagtgtttcaataaggaa 1571
          |||||||||||||||||||| ||| ||||||||||||||||||||||||||| | |||||||
Db  54682 GTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCGCCTCTTCTAGGTATTTTGGAATGTA 54623

Qy  1572  taact 1576
          |||||
Db  54622 CAATT 54618
```

```
Query Match     78.7%;  Score 187.8;  DB 9;  Length 97037;
Best Local Similarity  78.7%;  Pred. No. 3.4e-31;
Matches 240;  Conservative  0;  Mismatches 57;  Indels  8;  Gaps 1;

Qy  1280  agggtgagggtctcgtcttacacctagctcactgcaacc 1339
          || ||| ||||||||||| ||||||||||||||||||||
Db  54922 AGAGTCTTGCTCTGTCATCCAAGCTGGAGTGCAGTGGCACGGATCTCAGCTCACTGCAAC 54863

Qy  1340  tctgcctcccaggttcaagcaattcctcctgcctcagcctcccgctagctgggactacag 1399
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               complement(774..1726)
repeat_unit    /rpt_family="L1"
               /evidence=experimental
               complement(1842..2131)
repeat_unit    /rpt_family="L1"
               2312..2465
               /evidence=experimental
repeat_unit    /rpt_family="L1PA11"
               /evidence=experimental
               complement(2468..2757)
repeat_unit    /rpt_family="MIR2"
               /evidence=experimental
               complement(3204..3325)
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               2774..3059
repeat_unit    /rpt_family="MLTIC"
               /evidence=experimental
               3384..3674
repeat_unit    /rpt_family="Alu-Sq"
               /evidence=experimental
               3794..4216
repeat_unit    /rpt_family="L1MA9"
               /evidence=experimental
               7502..7792
repeat_unit    /rpt_family="Alu-J"
               /evidence=experimental
               7965..8072
repeat_unit    /rpt_family="L1MA2"
               /evidence=experimental
               8073..8360
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               8379..8535
repeat_unit    /rpt_family="L1MA5"
               /evidence=experimental
               8540..8743
repeat_unit    /rpt_family="L1"
               /evidence=experimental
               8744..8869
repeat_unit    /rpt_family="Alu-J"
               /evidence=experimental
               complement(9771..10055)
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               10292..10780
repeat_unit    /rpt_family="L1"
               /evidence=experimental
               11403..11692
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               12097..12459
repeat_unit    /rpt_family="MST"
               /evidence=experimental
               complement(12492..12701)
repeat_unit    /rpt_family="Alu-J or an Alu-S"
               /evidence=experimental
               complement(12730..13020)
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               13708..13997
repeat_unit    /rpt_family="Alu-Sq"
               /evidence=experimental
               complement(14402..14541)
repeat_unit    /rpt_family="MIR2"
               /evidence=experimental
               14546..14836
repeat_unit    /rpt_family="Alu-Sx"
               15348..15628
               /evidence=experimental
               /rpt_family="Alu-Sx"

repeat_unit    /evidence=experimental
               complement(15805..16906)
repeat_unit    /rpt_family="LTR12"
               /evidence=experimental
               17588..17930
repeat_unit    /rpt_family="MER44"
               /evidence=experimental
               complement(20000..20276)
repeat_unit    /rpt_family="Alu-Sb0 or an Alu-Sb1"
               /evidence=experimental
               20353..20855
repeat_unit    /rpt_family="L1ME3A"
               /evidence=experimental
               20900..20986
repeat_unit    /rpt_family="MSTA"
               /evidence=experimental
               20988..21276
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               21342..21723
repeat_unit    /rpt_family="MIR2"
               /evidence=experimental
               21841..22129
repeat_unit    /rpt_family="Alu-Sp"
               /evidence=experimental
               22150..22445
repeat_unit    /rpt_family="Alu-J"
               /evidence=experimental
               complement(22542..22789)
repeat_unit    /rpt_family="THE1b"
               /evidence=experimental
               complement(23235..23377)
repeat_unit    /rpt_family="MIR2"
               /evidence=experimental
               complement(23513..23798)
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               24296..24578
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               24808..24944
repeat_unit    /rpt_family="Alu-Sq"
               /evidence=experimental
               24945..25234
repeat_unit    /rpt_family="Alu-Sb2"
               /evidence=experimental
               25236..25393
repeat_unit    /rpt_family="Alu-Sc"
               /evidence=experimental
               complement(25486..26773)
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               26570..26859
repeat_region  /rpt_family="CpG Island"
               /evidence=experimental
               26801..27091
repeat_unit    /rpt_family="Alu-Sx"
               /evidence=experimental
               27866..27986
repeat_unit    /rpt_family="MIR"
               /evidence=experimental
               complement(28015..28185)
repeat_unit    /rpt_family="MIR"
               /evidence=experimental
               28550..29215
repeat_region  /rpt_family="CpG Island"
               /evidence=experimental
               31026..31323
repeat_unit    /rpt_family="MLR"
               /evidence=experimental
               complement(32698..32904)
repeat_unit    /rpt_family="Alu-Sc"
               /evidence=experimental
```

```
repeat_unit  33130..33399
             /rpt_family="MER21"

Query Match    3.5%;  Score 187.8;  DB 67;  Length 135038;
Best Local Similarity  78.7%;  Pred. No. 3.5e-31;
Matches 240;  Conservative 0;  Mismatches 57;  Indels 8;  Gaps 1;

Qy  1280 agggtgagggtctgtgtcttacacctacctgtatgctctcacctgagctcactgcaacc 1339
         ||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  8349 AGAGTCTTGCTGTGTCATCCAAGCTGGAGTGCAGTGGCACGAGATCTCAGCTCACTCCAACC 8290

Qy  1340 tctgcctcccaggttcaagcaaattcctcctgtcctcagcctcccgcgtagctgggactacag 1399
         |||||||||| |||||||||||||| ||||||| |||||||||||||| ||||||||||||||
Db  8289 TCTGCCTCCCTCGGGTTCAAGCAAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGACTACAG 8230

Qy  1400 gcg---------cacgcccggctaattttttgtatttttagtagagatggggtttcaccata 1451
         |||         ||||||| ||||||||| ||||||||||||||||||||||||||||||||
Db  8229 GCGCAATGCCACCACGCCCAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATA 8170

Qy  1452 ttagcccggcctggtcttgaactcctgacctcaggtgatccacccaccctcagcctcccctaa 1511
         || ||||||| ||||||||||||||||||||||| |||||||||| |||| |||||| |
Db  8169 TTGGTGCAGGCTGGTCTTGAACTCCTGACCTCAGATGATCCACCCTGCCTCGGCCTCCCAAA 8110

Qy  1512 gtgtctgggattacaggcatgagtcaccgcgcccggccaagggtcagtgttaataaggaa 1571
         ||| ||||||||||||| |||| |||||||||||||| |
Db  8109 GTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCCCTCTTCTAGGTATTTTGGAATGTA 8050

Qy  1572 taact 1576
         ||  |
Db  8049 CAATT 8045
```

```
RESULT 12
HS821D11
HS821D11/c
LOCUS       HS821D11    76727 bp   DNA          PRI    12-DEC-1999
DEFINITION  Human DNA sequence from clone RP5-821D11 on chromosome 22q12.3-13.1
            Contains three partial unknown genes, one downstream of a predicted
            CpG Island, and the first coding exon of the SREBF2 gene for Sterol
            Regulatory Element Binding Transcription Factor 2 downstream of a
            predicted CpG Island, ESTs, STSs, GSSs and genomic marker D22S1157,
            complete sequence.
ACCESSION   AL021453
VERSION     AL021453.1  GI:3413288
KEYWORDS    HTG; D22S1157; SREBF2; Sterol Regulatory Element Binding
            Transcription Factor 2.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76727)
AUTHORS     Bird,C.
TITLE       Direct Submission
JOURNAL     Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 12, 1998 this sequence version replaced gi:3355590.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP5-821D11 is
            from the library RPCI-5 constructed at the Roswell Park Cancer
            Institute by the group of Pieter de Jong. For further details see
            http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP5-821D11 It may be shorter because we only sequence overlapping
            sections once, or longer because we arrange for a small overlap
            between neighbouring submissions.
            The true left end of clone RP5-821D11 is at 1 in this sequence. The
            true left end of clone CTA-250D10 is at 76628 in this sequence. The
            true right end of clone CTA-109G6 is at 42082 in this sequence. The
            start of this sequence overlaps with sequence 299716 The end of
            this sequence overlaps with sequence 283840.
FEATURES             Location/Qualifiers
     source          1..76727
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q12.3-13.1"
                     /clone="RP5-821D11"
                     /clone_lib="RPCI-5"
     repeat_region   350..434
                     /note="L2 repeat: matches 2490..2537 of consensus"
     repeat_region   435..743
                     /note="AluJb repeat: matches 8..312 of consensus"
     variation       446..448
                     /note="clone CTA-109G6
                     tct in this entry
                     substitution"
     repeat_region   744..840
                     /note="L2 repeat: matches 2381..2490 of consensus"
     variation       ...
                     /gene="dJ821D11.1"
                     /note="clone CTA-109G6
                     tgt in this entry
                     /replace="tct"
                     substitution"
     mRNA            join(1230..1711,13695..13736)
                     /gene="dJ821D11.1"
                     /note="match: ESTs: Em:AA419437"
                     /note="other possible startcodon at 1583
                     this gene and dJ821D11.2 could be part of one gene"
     gene            1230..13736
                     /gene="dJ821D11.1"
                     /product="dJ821D11.1 (PUTATIVE protein)"
     gene            1314..1316
                     /gene="dJ821D11.1"
     variation       1314..1316
                     /gene="dJ821D11.1"
                     /note="clone CTA-109G6
     CDS             join(1601..1711,13695..>13736)
                     /gene="dJ821D11.1"
                     /note="match: proteins: Tr:O95505"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="dJ821D11.1 (PUTATIVE protein)"
                     /protein_id="CAA16279.1"
                     /db_xref="GI:4200030"
                     /db_xref="SPTREMBL:O95505"
                     /translation="MLATLAKADSPRTALLCSAWLLTASFSAQQHKGSLQKDPLLSQ
                     ACVGCLE"
     variation       1760..1762
                     /gene="dJ821D11.1"
                     /note="clone CTA-109G6
                     gag in this entry
                     /replace="ggg"
                     substitution"
     repeat_region   1773..1923
                     /note="MIR repeat: matches 47..201 of consensus"
     variation       1901..1903
                     /gene="dJ821D11.1"
                     /note="clone CTA-109G6
```

repeat_region   tca in this entry
                 substitution"
                 2010. .2209
                 /replace="tta"

variation   /note="L2 repeat: matches 2250. .2453 of consensus"
                 2302. .2304
                 gga in this entry
                 substitution"
                 3042. .3152
                 /replace="gaa"

repeat_region   /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 3162. .3457
                 /note="L1MC1 repeat: matches 6231. .6332 of consensus"

repeat_region   3458. .3551
                 /note="AluSp repeat: matches 1. .297 of consensus"

repeat_region   3552. .3848
                 /note="MIR repeat: matches 74. .148 of consensus"

repeat_region   3849. .3887
                 /note="AluJb repeat: matches 1. .299 of consensus"

variation   /note="MIR repeat: matches 32. .74 of consensus"
                 4753. .4754
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 tt in this entry
                 deletion"

repeat_region   /replace="ttgagt"
                 4772. .4790
                 /note="L1M4 repeat: matches 3167. .3185 of consensus"

repeat_region   4791. .5099
                 /note="AluSq repeat: matches 1. .312 of consensus"

variation   4791. .4792
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 tt in this entry
                 deletion"
                 /replace="ttttattat"

repeat_region   5100. .5116
                 /note="L1M4 repeat: matches 3151. .3167 of consensus"

variation   5107. .5112
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 tt in this entry
                 insertion"
                 /replace="tt"

repeat_region   5117. .5427
                 /note="AluYb8 repeat: matches 1. .318 of consensus"

variation   5359. .5361
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 att in this entry
                 substitution"
                 /replace="act"

repeat_region   5430. .5729
                 /note="AluSp repeat: matches 1. .300 of consensus"

repeat_region   5730. .5764
                 /note="L1M4 repeat: matches 3116. .3151 of consensus"

variation   5761. .5762
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 att in this entry
                 deletion"
                 /replace="att"

repeat_region   5765. .5075
                 /note="AluY repeat: matches 1. .311 of consensus"

variation   5772. .5773
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 tt in this entry
                 deletion"
                 /replace="tat"

variation   5778. .5779
                 /gene="dJ821D11.1"

---

variation   /note="clone CTA-109G6
                 tt in this entry
                 deletion"
                 /replace="ttat"
                 5780. .5782
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 tat in this entry
                 substitution"
                 /replace="tt"
                 5799. .5801
                 /gene="dJ821D11.1"
                 /note="clone CTA-109G6
                 cgg in this entry
                 substitution"

Query Match   3.5%;   Score 185.8;   DB 65;   Length 76727;
Best Local Similarity 78.5%;   Pred. No. 9.3e-31;
Matches 238; Conservative   0; Mismatches 57; Indels   8; Gaps 1;

```
Qy   1280 agggtgaggtctgtgtctacacctacctgtatgctctacacctgagctcactgcaacc 1339
          ||||||  |||||  ||  ||  |  |  |  |  || |||||||| |||||||||||||  ||
Db  57418 AGAGTCTCATTCTGTTTGCCTAGGCGTGGAGTGCAGTGGCATGAGTCTCAGCTTACTGCAACC 57359

Qy   1340 tctgcctcccaagtccaagacaattctcttcttgttccagcctccccccgttacgtggggactacag 1399
          |||||||||||  || |||| | ||||| |  |  ||| | |||||||||  || ||
Db  57358 TCTGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCGTCTCCCGAGGTTCAGCGTTGCTGGGACTATAG 57299

Qy   1400 gcg----------cacgcccggctaattttgtattgttagtagagatggggtttcaccata 1451
          |||          ||||||||||||||||| |||||| |||||||||||||||||||||||
Db  57298 GCGGTACCACCATGCCCGGCTAATTTTTGTATTTTAGTAGAGATGGGGTTTCACCATA 57239

Qy   1452 ttagccggcctggtcttgaactcctgacctcaggtgatccacccacccggccttcctaaa 1511
          || || | |||||||||||||||||||||||||||||||| |||||| ||||| || |||
Db  57238 TTGGCTAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCGCCCACCTCGGCCTCCCAAA 57179

Qy   1512 gtgctgggattacaggcgtgagtcacccgccgccggccaggggtcaagtgtttttattaaggaa 1571
          |||||||||||||||||  | ||| ||| | ||  ||||  |||| |   |||| ||| |  ||
Db  57178 GTGCTGGGATTACAGGCCATGAGCCACCACTGTCCCCGGCCCAAAAATCTGGGTTATTTCAAATAA 57119

Qy   1572 taa 1574
          |||
Db  57118 AAA 57116
```

RESULT 13
AC026395/c

LOCUS   AC026395   152044 bp   DNA   HTG   15-JUN-2000
DEFINITION Homo sapiens chromosome 10 clone RP11-45D20, WORKING DRAFT
            SEQUENCE, 44 unordered pieces.
ACCESSION   AC026395
VERSION   AC026395.3 GI:8567738
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 152044)
  AUTHORS   Smith,D.R.
  TITLE   Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 152044)
  AUTHORS   Smith,D.R.
  TITLE   Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT   On Jun 15, 2000 this sequence version replaced gi:7330305.
            -------------- Genome Center
                Center: Genome Therapeutics Corporation
                Center code: GTC
                Web site: http://www.genomecorp.com/
                Contact: gtc-seqcenter@genomecorp.com

```
                        --------  Project Information
                        --------  Summary Statistics
             Center project name: hg344
             Sequencing vector: N/A
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 990315
             Consensus quality: 126801 bases at least Q40
             Consensus quality: 138138 bases at least Q30
             Consensus quality: 140422 bases at least Q20
             Insert size: 147744; sum-of-contigs
             Quality coverage: 3.2x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
*      1006: contig of 1006 bp in length
*      1006: gap of unknown length
*      1007: contig of 1369 bp in length
*      2475: gap of unknown length
*      2476: contig of 1352 bp in length
*      3927: gap of unknown length
*      3928: contig of 1306 bp in length
*      5333: gap of unknown length
*      5334: contig of 1131 bp in length
*      5434: gap of unknown length
*      6564: contig of 1131 bp in length
*      6565: gap of unknown length
*      6664: contig of 1206 bp in length
*      7870: gap of unknown length
*      7871: contig of 1414 bp in length
*      9384: gap of unknown length
*      9385: contig of 1307 bp in length
*      9485: gap of unknown length
*     10791: contig of 1292 bp in length
*     10792: gap of unknown length
*     10892: contig of 2119 bp in length
*     13010: gap of unknown length
*     13011: contig of 1269 bp in length
*     13110: gap of unknown length
*     13111: contig of 1507 bp in length
*     14617: gap of unknown length
*     14618: contig of 1500 bp in length
*     14718: gap of unknown length
*     16217: contig of 1789 bp in length
*     16218: gap of unknown length
*     16318: contig of 1789 bp in length
*     18106: gap of unknown length
*     18107: contig of 1292 bp in length
*     18207: gap of unknown length
*     19498: contig of 1307 bp in length
*     19499: gap of unknown length
*     19599: contig of 1269 bp in length
*     20867: gap of unknown length
*     20868: contig of 1236 bp in length
*     20967: gap of unknown length
*     20968: contig of 1236 bp in length
*     22203: gap of unknown length
*     22204: contig of 1795 bp in length
*     23998: gap of unknown length
*     24099: contig of 1807 bp in length
*     24198: gap of unknown length
*     24199: contig of 1807 bp in length
*     26005: gap of unknown length
*     26006: contig of 1336 bp in length
*     26105: gap of unknown length
*     26106: contig of 1336 bp in length
*     27441: gap of unknown length
*     27442: contig of 1496 bp in length
*     27541: gap of unknown length
*     27542: contig of 1496 bp in length
*     29037: gap of unknown length
*     29038: contig of 1523 bp in length
*     29137: gap of unknown length
*     29138: contig of 1523 bp in length
*     30660: gap of unknown length
*     30661: contig of 1995 bp in length
*     30760: gap of unknown length
*     30761: contig of 1995 bp in length
*     32855: gap of unknown length
*     32856: contig of 1628 bp in length
*     34483: gap of unknown length
*     34484: contig of 2995 bp in length
*     37578: gap of unknown length
*     37579: contig of 2021 bp in length
*     39699: gap of unknown length
*     39700: contig of 2612 bp in length
*     42411: gap of unknown length
*     42412: contig of 3127 bp in length
*     45638: gap of unknown length
*     45639: contig of 4001 bp in length
*     49739: contig of 4001 bp in length
```

```
*     49740: gap of unknown length
*     49840: contig of 4437 bp in length
*     54276: gap of unknown length
*     54277: contig of 4675 bp in length
*     54376: gap of unknown length
*     54377: contig of 4675 bp in length
*     59051: gap of unknown length
*     59052: contig of 5475 bp in length
*     59151: gap of unknown length
*     59152: contig of 5475 bp in length
*     64626: gap of unknown length
*     64627: contig of 3252 bp in length
*     64726: gap of unknown length
*     64727: contig of 3252 bp in length
*     67988: gap of unknown length
*     67989: contig of 4406 bp in length
*     68088: gap of unknown length
*     68089: contig of 4406 bp in length
*     72494: gap of unknown length
*     72495: contig of 3049 bp in length
*     72594: gap of unknown length
*     72595: contig of 3049 bp in length
*     75643: gap of unknown length
*     75644: contig of 3767 bp in length
*     75743: gap of unknown length
*     75744: contig of 3767 bp in length
*     79510: gap of unknown length
*     79511: contig of 5054 bp in length
*     79610: gap of unknown length
*     79611: contig of 5054 bp in length
*     84664: gap of unknown length
*     84665: contig of 5868 bp in length
*     84764: gap of unknown length
*     84765: contig of 5868 bp in length
*     90513: gap of unknown length
*     90514: contig of 5750 bp in length
*     90613: gap of unknown length
*     90614: contig of 5750 bp in length
*     96184: gap of unknown length
*     96185: contig of 5571 bp in length
*     96285: gap of unknown length
*     96286: contig of 5763 bp in length
*    102048: gap of unknown length
*    102049: contig of 3767 bp in length
*    102148: gap of unknown length
*    102149: contig of 6213 bp in length
*    108361: gap of unknown length
*    108362: contig of 6213 bp in length
*    108461: gap of unknown length
*    108462: contig of 5868 bp in length
*    114329: gap of unknown length
*    114330: contig of 6215 bp in length
*    114429: gap of unknown length
*    114430: contig of 6215 bp in length
*    120644: gap of unknown length
*    120645: contig of 7281 bp in length
*    120744: gap of unknown length
*    120745: contig of 7281 bp in length
*    128025: gap of unknown length
*    128026: contig of 10663 bp in length
*    128125: gap of unknown length
*    128126: contig of 10663 bp in length
*    138788: gap of unknown length
*    138789: contig of 13156 bp in length
*    138888: gap of unknown length
*    138889: contig of 13156 bp in length
*    152004: contig of 13156 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..152044
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-45D20"
                     /clone_lib="RPCI-11"

BASE COUNT     40176 a     33928 c     33229 g     40362 t     4349 others
ORIGIN
```

```
Qy     1321  acctgagctcactgaacccctcgcctcccaggttcaagcaattcctcgtcctcagcctcc  1380
              |||||||||||||||||||||||||| |||||| |||||| ||||||||||||||||||||
Db    30398  ATCTCGGCTCACTGCAACCTCCACTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCC   30339

Qy     1381  cgcgtagctgggactacaggc----------gcacgcccggctaattctttgtattgttagta  1432
              |||||||||||||||||||||           |||||||||||||||||||||||||||||||
Db    30338  TGAGTAGCTGGGATTACAGGCACCACGCCCAGCTAATTTTTGTATTTTTAGTA   30279

Qy     1433  gagatagggggtttcaccacatattagcccggctcggtcttgaacctcctcaggtgatcca  1492
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    30278  GAGATGGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCA   30219

Qy     1493  cccacctcagcctcctaaagtgctggggattacagaggcatgagtcaccgccggcgccaagg  1552
              ||| ||||||||||||||||||||| ||||||| ||||| |||| |||||||||| ||||||
Db    30218  CCCGCCTTAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACTGGGCCAATT   30159

Qy     1553  gtcagtgtttaataagggaataacttgaatggtttactaaccacag  1599
              |||||||||||||||||||||||||||||||||||||||||||||||
Db    30158  TTAATTTTTTTTTAATGAAAAAATGTGGATATTTTAAAAAAAACACAG   30112
```

```
Query Match          3.5%;     Score 185.8;  DB 79;   Length 152044;
Best Local Similarity  80.8%;  Pred. No. 9.6e-31;
Matches  232;  Conservative    0;  Mismatches   47;  Indels    8;  Gaps   1;
```

```
LOCUS       AC025947      157057 bp    DNA             HTG      10-JUN-2000
ACC025947/c    AC025947    157057 bp
RESULT  14
```

```
DEFINITION   Homo sapiens chromosome 10 clone RP11-78A18, WORKING
             DRAFT SEQUENCE, 26 unordered pieces.
ACCESSION    AC025947
VERSION      AC025947.3  GI:8439851
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 157057)
  AUTHORS    Smith,D.R.
  TITLE      Genome Therapeutics Corporation Sequencing Center: Human Genome
             Sequence Data
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 157057)
  AUTHORS    Smith,D.R.
  TITLE      Direct Submission
  JOURNAL    Submitted (18-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
             Street, Waltham, MA 02453, USA
COMMENT      On Jun 10, 2000 this sequence version replaced gi:7528340.
             -------------- Genome Therapeutics Corporation
             Center: Genome Therapeutics Corporation
             Center code: GTC
             Web site: http://www.genomecorp.com/
             Contact: gtc-seqcenter@genomecorp.com
             --------------- Project Information
             Center project name: hg307
             --------------- Summary Statistics
             Sequencing vector: N/A
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 990315
             Consensus quality: 140586 bases at least Q40
             Consensus quality: 148596 bases at least Q30
             Consensus quality: 150061 bases at least Q20
             Insert size: 154557; sum-of-contigs
             Quality coverage: 3.7x in Q20 bases; sum-of-contigs
             --------------
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 26 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *
           *     1     1544:  contig of 1544  bp in length
           *  1545     1644:  gap of unknown  length
           *  1645     3218:  contig of 1574  bp in length
           *  3219     3318:  gap of unknown  length
           *  3319     5022:  contig of 1704  bp in length
           *  5023     5122:  gap of unknown  length
           *  5123     6553:  contig of 1431  bp in length
           *  6554     6653:  gap of unknown  length
           *  6654     8651:  contig of 1998  bp in length
           *  8652     8751:  gap of unknown  length
           *  8752    10843:  contig of 2092  bp in length
           * 10844    10943:  gap of unknown  length
           * 10944    13772:  contig of 2829  bp in length
           * 13773    13872:  gap of unknown  length
           * 13873    15825:  contig of 1953  bp in length
           * 15826    15925:  gap of unknown  length
           * 15926    19591:  contig of 3666  bp in length
           * 19592    19691:  gap of unknown  length
           * 19692    22350:  contig of 2659  bp in length
           * 22351    22450:  gap of unknown  length
           * 22451    25948:  contig of 3498  bp in length
           * 25949    26048:  gap of unknown  length
           * 26049    29641:  contig of 3593  bp in length
           * 29642    29741:  gap of unknown  length
           * 29742    34105:  contig of 4364  bp in length
           * 34106    34205:  gap of unknown  length
           * 34206    38484:  contig of 4279  bp in length
           * 38485    38584:  gap of unknown  length
           * 38585    41992:  contig of 3408  bp in length
           * 41993    42092:  gap of unknown  length
           * 42093    47920:  contig of 5828  bp in length
           * 47921    48020:  gap of unknown  length
           * 48021    54416:  contig of 6396  bp in length
           * 54417    54516:  gap of unknown  length
           * 54517    60194:  contig of 5678  bp in length
           * 60195    60294:  gap of unknown  length
           * 60295    67639:  contig of 7345  bp in length
           * 67640    67739:  gap of unknown  length
           * 67740    74790:  contig of 7051  bp in length
           * 74791    74890:  gap of unknown  length
           * 74891    85531:  contig of 10741 bp in length
           * 85532    85731:  gap of unknown  length
           * 85732    93940:  contig of 8209  bp in length
           * 93941    94040:  gap of unknown  length
           * 94041   104191:  contig of 10151 bp in length
           * 104192  104291:  gap of unknown  length
           * 104292  116387:  contig of 12096 bp in length
           * 116388  116487:  gap of unknown  length
           * 116488  135857:  contig of 19370 bp in length
           * 135858  135957:  gap of unknown  length
           * 135958  157057:  contig of 21100 bp in length.
FEATURES             Location/Qualifiers
     source          1..157057
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="Chromosome 10"
                     /clone="RP11-78A18"
                     /clone_lib="RPCI-11"
BASE COUNT    45178 a   32550 c   32372 g   44448 t    2509 others
ORIGIN

Query Match     3.5%;  Score 185.8;  DB 78;  Length 157057;
Best Local Similarity 80.8%;
Matches  232;  Conservative   0;  Mismatches  47;  Indels   8;  Gaps   1;

Qy   1321 acctgagctcactgcaacctctgccccaggttcaagcaattcctccctgtcctcagcctcc 1380
          ||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Db 109551 ATCTCGGCTCACTGCAACCTCCCACCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCC 109492

Qy   1381 cgcctagctggggactacaggc---------gcacgcccggctaattttttgtattgttagta 1432
          |||||||||||||||||||||           ||||||||||||||||||| ||||||||||
Db 109491 TGAGTAGCTGGGATTACAGGCGCACCCACTACCACGCCCAGCTAATTTTGTATTTTTAGTA 109432

Qy   1433 gagatgggggttcaccacatattagcccggtcttgaactcctgacctcaggtgatcca 1492
          |||||||||||||||||||||| || |||| |||||||||||||||||||||||||||
Db 109431 GAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCA 109372

Qy   1493 cccacctcagcctcctaaagtgctgggattacaggcatgagtcaccgcgcccggcccaagg 1552
          ||||||||||||||| ||||||||||||||||||||||||| ||| || ||||||||||||
Db 109371 CCCGCCTTAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCGCACTGCACCTGGCC 109312

Qy   1553 gtcagtgttaataaggaataacttgaatggttttactaaccaacag 1599
          ||||||||||||||||||||||||||||||||||||||||||||||
Db 109311 TTAATTTTTTTTAATGAAAAAAATGTGGATATTTTTAAAAAAACACAG 109265
```

RESULT 15
```
LOCUS       AC015488
ACCESSION   AC015488
LOCUS       AC015488    161499 bp    DNA           HTG      16-MAR-2000
DEFINITION  Homo sapiens clone RP11-20F6, WORKING DRAFT SEQUENCE, 16 unordered
            pieces.
ACCESSION   AC015488
VERSION     AC015488.4  GI:7107977
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161499)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
```

```
TITLE      Homo sapiens, clone RP11-20F6
JOURNAL    Unpublished
REFERENCE  2  (bases 1 to 161499)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
           Baldwin,J., Barna,N., Beckerly,R., Boguslavkty,L., Boukhgalter,B.,
           Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
           Cooke,P., DeArellano,K., Dewar,K., Donelan,L., Doyle,M.,
           Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
           Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
           Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
           Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
           McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
           Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
           Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
           Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
           Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
           Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
           Direct Submission
           Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
           On Feb 28, 2000 this sequence version replaced gi:6715933.
           All repeats were identified using RepeatMasker:
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           Smit,A.F.A. & Green,P. (1996-1997)
           -------------------------------------------------------
           Center: Whitehead Institute/ MIT Center for Genome Research
           Center code: WIBR   Genome Center
           Web site: http://www-seq.wi.mit.edu
           Contact: sequence_submissions@genome.wi.mit.edu
           -------------------------------------------------------
           Project Information
           Center project name: L3895
           Center clone name: 20_F_6
           -------------------------------------------------------
           Summary Statistics
           Sequencing vector: M13; M77815; 100% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads
           Assembly program: Phrap; version 0.960731
           Consensus quality: 152793 bases at least Q40
           Consensus quality: 156653 bases at least Q30
           Consensus quality: 158136 bases at least Q20
           Insert size: 160000; agarose-fp
           Insert size: 159999; sum-of-contigs
           Quality coverage: 5.9 in Q20 bases; agarose-fp
           Quality coverage: 5.9 in Q20 bases; sum-of-contigs
           -------------------------------------------------------

     * NOTE: This is a 'working draft' sequence. It currently
     * consists of 16 contigs. The true order of the pieces
     * is not known and their order in this sequence record is
     * arbitrary. Gaps between the contigs are represented as
     * runs of N, but the exact sizes of the gaps are unknown.
     * This record will be updated with the finished sequence
     * as soon as it is available and the accession number will
     * be preserved.
     *
     *         1  2206: contig of 2206 bp in length
     *      2207  2306: gap of 100 bp
     *      2307  6910: contig of 4604 bp in length
     *      6911  7010: gap of 100 bp
     *      7011  9818: contig of 2808 bp in length
     *      9819  9918: gap of 100 bp
     *      9919 13394: contig of 3476 bp in length
     *     13395 13494: gap of 100 bp
     *     13495 17122: contig of 3628 bp in length
     *     17123 17222: gap of 100 bp
     *     17223 19752: contig of 2530 bp in length
     *     19753 19852: gap of 100 bp
     *     19853 23958: contig of 4106 bp in length
     *     23959 24058: gap of 100 bp
     *     24059 29684: contig of 5626 bp in length
     *     29685 29784: gap of 100 bp
     *     29785 33951: contig of 4167 bp in length
     *     33952 34051: gap of 100 bp
     *     34052 39841: contig of 5790 bp in length
     *     39842 39941: gap of 100 bp
     *     39942 48297: contig of 8356 bp in length
```

```
     *     48298 48397: gap of          100 bp in length
     *     48398 56898: contig of 8501 bp in length
     *     56899 56998: gap of          100 bp
     *     56998 65924: contig of 8926 bp in length
     *     65925 66024: gap of          100 bp
     *     66025 81114: contig of 15090 bp in length
     *     81115 81214: gap of          100 bp
     *     81215 112500: contig of 31286 bp in length
     *    112501 112600: gap of          100 bp in length
     *    112601 161499: contig of 48899 bp in length.

FEATURES             Location/Qualifiers
     source          1..161499
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-20F6"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..2206
                     /note="assembly_fragment"
     misc_feature    2307..6910
                     /note="assembly_fragment"
     misc_feature    7011..9818
                     /note="assembly_fragment"
     misc_feature    9919..13394
                     /note="assembly_fragment"
     misc_feature    13495..17122
                     /note="assembly_fragment"
     misc_feature    17223..19752
                     /note="assembly_fragment"
     misc_feature    19853..23958
                     /note="assembly_fragment"
                     /clone_end:T7
                     /vector_side:right"
     misc_feature    24059..29684
                     /note="assembly_fragment"
     misc_feature    29785..33951
                     /note="assembly_fragment"
     misc_feature    34052..39841
                     /note="assembly_fragment"
     misc_feature    39942..48297
                     /note="assembly_fragment"
     misc_feature    48398..56898
                     /note="assembly_fragment"
     misc_feature    56998..65924
                     /note="assembly_fragment"
     misc_feature    66025..81114
                     /note="assembly_fragment"
                     /clone_end:SP6
                     /vector_side:left"
     misc_feature    81215..112500
                     /note="assembly_fragment"
     misc_feature    112601..161499
                     /note="assembly_fragment"

BASE COUNT    44611 a    37543 c    36473 g    41370 t    1502 others
ORIGIN
```

```
Query Match     3.5%;  Score 185.8;  DB 72;  Length 161499;
Best Local Similarity   77.1%;  Pred. No. 9.7e-31;
Matches  226;  Conservative   0;  Mismatches  67;  Indels   0;  Gaps   0;

Qy    1283 gtgagggtctgtgtcttacacctacctgtatgctctacacctgagctcactgcaacctct 1342
           || |||||| |||| || ||||||||||||  ||||||||||||||||||||||||||||
Db   16769 GTCTTGCTCTGTCACCCAGGCTGGAGGGCAGTGGCACCATCTCGGCTCACTGCAACCTCT 16828

Qy    1343 gcctccaggtcaagcaaattcctcccctgtctcagcctccgcgtagctgggactacaggcg 1402
           |||||||||||||||| |||||||||||||||||||||||||| |||| |||| || ||
Db   16829 GCCTCCAGATTCAAGCAAATTCTCTCCCCTCAGCCCTCCTGAGTACCTGGGGTTATAGGCA 16888

Qy    1403 cacgccgcggctaatttttgtattgttagtagaagatggggtttcaccattagcccggct 1462
           |||||||| ||| ||||||||||||||||||||  | |||| |||||||||  |  ||||
Db   16889 CATGCCACGGCTAATTTTTGTGTTTTTAGTAGAGATGGGGTTTTTGCCATGTTGGCCAGGCT 16948

Qy    1463 ggtcttgaactcctgacctcaggtgatccaccacctcagcctcctaaagtgctgggatt 1522
```

```
Db  16949  GGTCTTGAACTTCCTGACCTCAGATGATCCGCCGCCTTCGGGCCTCCCAAAGTGCTGGGATT  17008
           ||||||||||||||||||||||| |||||| ||||||| |||| ||| |||||||||||||||
Qy   1523  acaggcatgagtcaccgccgcccggccaaagggtcagtgtttaataaggaataac  1575
           ||||| ||||||||||||||||||||||||||-|-||||-|-|||||-| |||
Db  17009  ACAGGTGTGAGCCACCGCGCCCAGCCTATTTTTAATTTTTATTTTTTATCAAC  17061
```

Search completed: December 3, 2000, 18:25:29
Job time: 43343 sec